nature

www.nature.com/nature

Vol 445 | Issue no. 7128 | 8 February 2007

Light at the end of the tunnel

An emphatic and clear status report on global warming opens the way for action — presenting new risks.

The release of the 2007 report of the Intergovernmental Panel on Climate Change (IPCC) last Friday marks an important milestone (see pages 578–585 and 595–598). Following the scientific consensus that has been apparent for some time, a solid political consensus that acknowledges the problem finally seems to be within reach. But achieving this outcome brings its own risks.

Until quite recently (perhaps even until last week), the general global narrative of the great climate-change debate has been deceptively straightforward. The climate-science community, together with the entire environmental movement and a broad alliance of opinion leaders ranging from Greenpeace and Ralph Nader to Senator John McCain and many US evangelical Christians, has been advocating meaningful action to curtail greenhouse-gas emissions. This requirement has been disputed by a collection of money-men and some isolated scientists, in alliance with the current president of the United States and a handful of like-minded ideologues such as Australia's prime minister John Howard.

The IPCC report, released in Paris, has served a useful purpose in removing the last ground from under the climate-change sceptics' feet, leaving them looking marooned and ridiculous. However, this predicament was already clear enough. Opinion in business circles, in particular, has moved on. A report released on 19 January by Citigroup, *Climatic Consequences* — the sort of eloquently written, big-picture stuff that the well-informed chief executive reads on a Sunday afternoon — states even more firmly than the IPCC that anthropogenic climate change is a fact that world governments are moving to confront. It leaves no question at all that large businesses need to get to grips with this situation — something that many of them are already doing.

Tough choices

So then, the enemy is vanquished and the victors can rejoice? Hardly. In fact, the pending retreat from the stage of the president of the United States and his allies leaves those who do acknowledge the severity of the problem facing an even greater challenge than before. The world now broadly accepts that we have a problem, if not a crisis. So what is to be done?

The policy choices that lie ahead are more daunting than political leaders (or the media) have thus far been ready to acknowledge. In a sense, twenty years of frustrating trench-warfare with the sceptics has prevented a rational discussion about what needs to be done from even taking place.

At present, the political response to the situation is, in large part, incongruous. We need to restrict emissions in the developed world, and some steps are being undertaken to do just that, chiefly through the much-maligned Kyoto Protocol. We need to develop clean energy sources, and these are being pushed ahead quite rapidly, although each one — nuclear power, biofuels, wind power and hydropower, for example — creates its own environmental battlefield. Steps are also being taken to build systems for large-scale carbon capture and storage, and to improve the efficiency with which energy is used (see pages 586–591).

The trouble is, none of this is even close to being sufficient to meet the challenge. Hybrid cars are being purchased (and often allow their lucky drivers special access to empty highway lanes). David Cameron, the leader of Britain's Conservative Party, has sought planning permission to erect a wind turbine in his back garden. And Pink Floyd and Pearl Jam have declared that their most recent world tours would be 'carbon neutral'. But we are all vaguely aware that all of this is nowhere near enough.

Economic sacrifice

Even the most progressive governments continue to put the issue of climate change on the back seat behind their fundamental commitment to strong economic growth, which is needed to ensure political survival (in developed countries) and to enable human dignity (in developing countries). So in a typical European nation, for example, governments are calling for strenuous emissions cuts while also planning airport expansions that anticipate a further tripling over the next twenty years of air travel — the fastest-growing source of emissions, and one not capped by the Kyoto Protocol.

The fundamental difficulty here is that it has been politically impossible to accept that fighting global warming may involve some economic sacrifice, at least while the sceptics were in the picture. As these are vanquished, it becomes possible — and indeed necessary — to start the discussion.

Similarly, it has been hard to talk about actions that need to be taken to mitigate the damage already certain to be caused by climate change and associated rises in the sea level, as such steps were regarded as a capitulation to those who just want to keep emitting greenhouse gases. This is no longer the case (see page 597). Mitigation, which can take many forms ranging from the Thames Barrier in London to the introduction of drought-resistant crop strains in the Sahel and the establishment of a proposed climate-change adaptation fund, needs to be squarely on the agenda, alongside emissions cuts.

"The IPCC report has served a useful purpose in removing the last ground from under the sceptics' feet, leaving them looking marooned and ridiculous."

A similar relaxation arises with regard to revised negotiations for the second stage of the Kyoto Protocol. There is a case for opening the second phase beyond a simple extension of the cap-and-trade proposals that made up the core of the first. US President George W. Bush will remain a participant in such negotiations until the end of 2008. But even before then, talks should include all the options open to a planet that is now ready, at last, to acknowledge the fix it is in. ■

Virgin offers big reward for capturing carbon

A US$25-million prize — one of the largest science prizes around — is on offer to the inventor of a device that will remove 'significant amounts' of carbon dioxide from the atmosphere.

"The winner must be able to demonstrate a commercially viable design which will result in the net removal of anthropogenic, atmospheric, greenhouse gases each year for at least ten years without countervailing harmful effects," state the rules of the Virgin Earth Challenge competition.

Among the judges are NASA climate scientist James Hansen; James Lovelock, inventor of the Gaia hypothesis; UK environmentalist and retired diplomat Crispin Tickell; and Australian conservationist and author Tim Flannery.

The competition is open for at least the next five years. British billionaire entrepreneur Richard Branson and former US vice-president Al Gore launched the prize on 9 February in London.

Historian to lead science forward at Harvard

Harvard University's new president, Drew Gilpin Faust, is a historian. But she is expected to continue many of the science-based initiatives that her predecessor, the controversial Larry Summers, put in place.

Plans are moving apace to create a science research complex in the Boston suburb of Allston — one of Summers' signature projects. He announced his resignation last February, in part over the furore about his comments on differences between men and women in science.

After her appointment on 11 February, Faust said that much remained to be done to address gender inequality in the sciences. She was the founding dean of the Radcliffe Institute for Advanced Study, a former women's college, which has put on science conferences on topics such as computational biology and tissue engineering. She also led two task forces looking at the status of women faculty at Harvard, including one on women in science and engineering, following Summers' controversial remarks.

Biologist Thomas Cech, head of the Howard Hughes Medical Institute in Chevy Chase, Maryland, was also a finalist for the president's job, but withdrew from the running last month.

K. SNIBBE/HARVARD

Pointing the way: Harvard boss Drew Gilpin Faust wants to address sexual inequality in science.

Horse genome to help human conditions

Twilight, the thoroughbred mare pictured here, is now the gold standard for all other horses: her DNA has been sequenced as the reference genome of the horse (*Equus caballus*).

The draft sequence, released on 7 February, should help the study of human conditions such as allergic disease, arthritis, exercise physiology and fertility. These studies will be aided by the fact that people have been breeding horses for 4,000 to 6,000 years and keeping close records of the bloodlines, allowing particular genes to be traced back through time.

D. ANTCZAK/CORNELL UNIV.

UK's Diamond synchrotron turns on the lights

Britain's Diamond synchrotron, a £260-million (US$505-million) device that is one of the country's largest pieces of new scientific infrastructure in decades, welcomed its first users late last month.

Diamond's initial users will study materials used in computer memories, a protein involved in cancer, and mineral samples from a meteorite that could shed light on conditions in the early Solar System. The facility currently has seven beamlines, producing X-rays with energies of between 100 and 20,000 electronvolts. It has funding to build another 15, which should all be online by 2011.

The debut of Diamond, situated near Oxford, marks the return of a world-class synchrotron to Britain: the first synchrotron was run in London in 1946. In 1993, government advisers warned that the country needed to build a new device to compete with systems elsewhere.

Catholic college sells land, but blocks stem-cell work

The University of Sydney in Australia will not conduct fetal stem-cell research in a new Aus$350-million (US$270 million) biomedical research centre planned on land obtained from a Roman Catholic college.

The 5 February compromise was reached after the university paid at least Aus$11 million to St John's College for land the Catholic school had controlled under a trust. St John's governing council sought the stem-cell restrictions even though the university had not planned such research at the new centre, which is expected to be completed by 2012.

University officials say they don't believe the agreement will create a precedent limiting stem-cell research at Australian public universities, such as Sydney.

Mars probe gets tunnel vision — in monochrome

The High-Resolution Imaging Science Experiment (HiRISE) on board the Mars Reconnaissance Orbiter — the newest and most powerful craft to arrive at the red planet — has lost its peripheral vision. And its colour vision is fading too.

Seven of HiRISE's 14 detectors are sending back spurious data, the mission team reports, and one of the four colour detectors has stopped working completely. This has led to only a 2% loss of signal so far, but the problem looks set to hit all of the detectors eventually.

"We do think it's a systematic problem for all of them," says Alfred McEwen, HiRISE's principal investigator, who is based at the University of Arizona in Tucson. "It's going to be a real irritant as it worsens."

Correction

The Editorial 'Light at the end of the tunnel' (*Nature* **445,** 567; 2007) should have referred to projects such as the Thames Barrier as 'adaptation' not 'mitigation'. In the related News story 'What we don't know about climate change' (*Nature* **445,** 580; 2007), the 2001 IPCC estimate for the range of sea-level rise should have read 9 to 88 centimetres, not millimetres.

Steady progress

Better budgets for biologists.

After years of rather disappointing funding news, things are looking up a little for US life scientists. On 31 January, Congress pledged to raise the 2007 budget of the National Institutes of Health (NIH) by $620 million to $29 billion, dispelling fears that the agency would be facing flat funding this year. But on 5 February, President George W. Bush called for 2008 funding at the biomedical research agency to be less than that (see page 572) — although Congress is likely to supplement it later in the year.

These developments together mark a modest success for the biomedical research lobby and for NIH director Elias Zerhouni. In the face of considerable scepticism, he argued that the NIH has made effective use of the doubling in funding that it received between 1998 and 2003. His push for 'translational medicine', with its emphasis on the transfer of ideas from the laboratory to the clinic, has helped to convince lawmakers that the agency is working hard to address America's healthcare needs.

The rapid expansion of the NIH was always likely to create a period of inflated expectations in biomedical research, during which researchers might feel short-changed. The funding boost was accompanied by a massive building boom at university medical centres and by a flood of postdocs to staff them. It was inevitable, after five years of 15% compound growth came to an end in 2003, that someone's expectations weren't going to be met. The best that the life sciences could hope for was that there would be only a short pause before Congress resumed its long-standing love affair with the NIH.

This pause may now be coming to an end. If that were so, it would be due to the efforts of Zerhouni and his staff, but also to scientists' direct lobbying on their own behalf. Organizations such as the Federation of American Societies for Experimental Biology (FASEB) and the American Society for Cell Biology (ASCB) have long urged researchers to become more involved in the political process, but have often felt that their calls go unheeded.

This time, scientists were listening. FASEB's Office of Public Affairs has announced that more than 15,000 scientists sent letters last summer urging public officials to reverse the slowdown in NIH funding, and more than 12,000 researchers have signed up to the organization's urgent-action e-mail list. At the ASCB meeting in December, a session on outreach and communication was packed, with standing-room only.

It would be unfortunate if this slightly better funding picture convinced researchers to drop their new-found level of political engagement. Scientists have plenty to do in their daily lives and may struggle to find time to write a letter to the newspaper, speak to a classroom at school, or even visit their Congressman. But in the end, such activities can pay dividends. ■

Welcome home

Italian and Spanish researchers returning from abroad deserve more support.

Most European nations bemoan an academic 'brain drain' that flows mostly in the direction of the United States. But the situation is not uniform. Germany, for example, is prone to grumble, but it has a well-funded, well-organized research infrastructure, and a regular supply of new academic positions. More than 80% of German scientists who go abroad for their PhDs or postdocs eventually return home.

Not all countries are so fortunate. When the infrastructure at home is poor, few migrants will return to the nest. Italy and Spain have both endured a prolonged haemorrhaging of talent. Having failed to invest properly in basic research for decades, both nations have recently attempted to plan a better future. In 2001, they each launched programmes specifically designed to entice well-trained scientists back home, by providing an attractive salary and research funding for up to five years.

Italy's 'Rientro dei cervelli' ('Brain gain') programme has so far brought back over 460 scientists; Spain's Ramon y Cajal scheme has attracted more than 2,000. The initial idea of both programmes was that researchers returning home would then be ideally placed to compete for local, permanent jobs. But thanks to bad planning, few of the returnees are as well-integrated as they had hoped. In the underfunded systems they found themselves in, few jobs became available.

The next step was therefore to encourage universities to create the required positions. Last year, Italy made an offer that, although modest, seemed to provide a solution. The government allocated €3 million (US$3.9 million) a year to pay for 95% of the salaries of returning scientists who were deemed worthy of a permanent position by their university departments.

But much of the money remains unspent. Many universities did not nominate candidates — partly because they see the 'Brain gain' programme as institutionalized queue-jumping, pushing returning scientists ahead of those who stayed at home waiting for jobs to show up. Most of the nominations were blocked on technical grounds by the similarly unsympathetic National Committee of Universities, which has to verify the eligibility of all candidate professors.

"Many universities see the 'Brain gain' programme as institutionalized queue-jumping, pushing returning scientists ahead of those who stayed at home."

Two years ago, the Spanish government made a similar offer, to fund the full costs of the first three years of 900 new academic posts. But uptake by universities has been patchy.

Full cooperation from the universities is essential if such schemes are to work. It would not have been easy — universities hate being told who to hire — but it should have been sought from the start.

Measures to bring in new talent are essential if Spain and Italy are to meet their scientific aspirations. It is now up to the universities to see the bigger picture and show more flexibility in supporting returning scientists. ■

RESEARCH HIGHLIGHTS

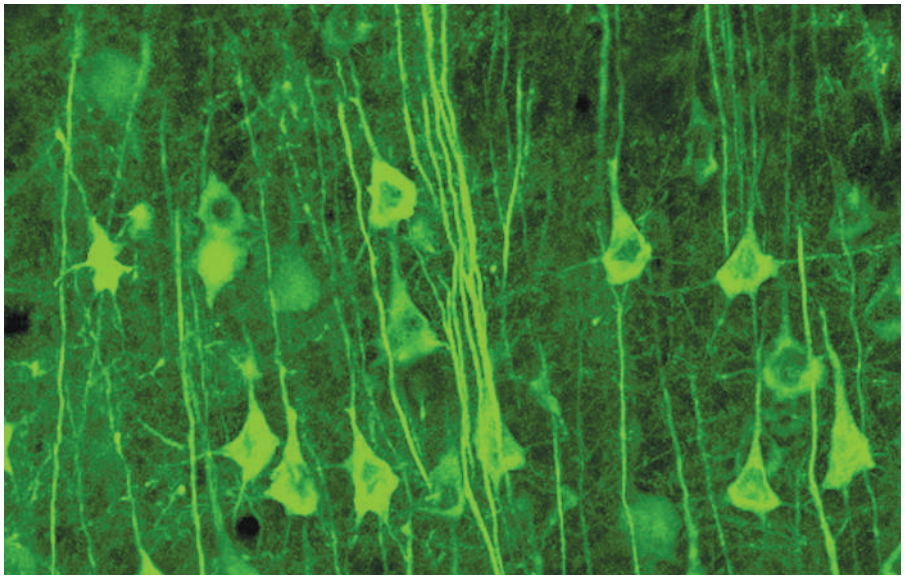
Neurons brought to light

Nature Meth. doi:10.1038/nmeth1009 (2007)

Researchers in Germany have developed a biosensor to image the activity of individual neurons (pictured) in the brains of living mice. The sensor detects the calcium ions involved in neuronal signalling.

Oliver Griesbeck from the Max Planck Institute for Neurobiology in Martinsried and his colleagues made the biosensor by attaching fluorescent proteins to a natural calcium-sensor protein, troponin C. The spectrum of light emitted by the biosensor changes when it binds to calcium.

Previous calcium sensors have performed poorly when engineered into mice, so the team modified the gene encoding this package to ensure that it would express a functional protein. The biosensor was sensitive enough for relatively low levels of electrical activity to be detected.



OPTOELECTRONICS

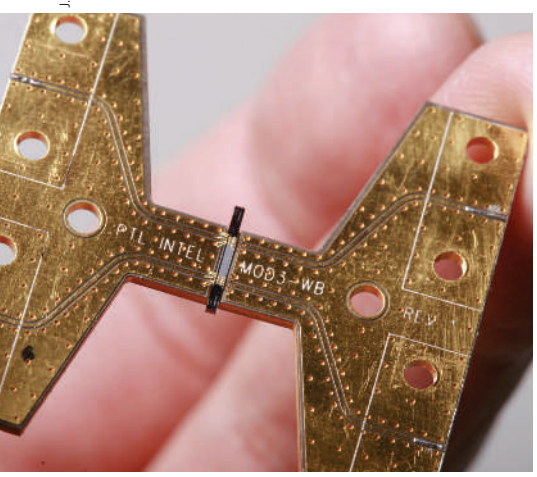
Speedy silicon

Optics Express **15,** 660–668 (2007)

Silicon is the material of choice for computer chips, but has traditionally lagged behind in optics, unable to compete with other materials for speed. That could soon change.

Ansheng Liu and his colleagues at Intel in Santa Clara, California, and in Jerusalem, Israel, have developed an optical device (pictured below) that can write up to 30 gigabits of data per second to a laser beam — three times faster than any previous silicon device. The optical modulator's speed stems in part from a design that transmits the electrical signal carrying the data and the light through a single channel, or 'waveguide'. The researchers claim they can tweak their device to reach data transfer speeds of up to 40 gigabits per second, which rivals the speeds of optical modulators made from more exotic materials.

J. TSENG/INTEL CORP.



NEUROBIOLOGY

Prion symptoms reversed

Neuron **53,** 325–335 (2007)

Early symptoms of the neurodegeneration caused by prion disease can be reversed in genetically engineered mice, report Giovanna Mallucci, of the Medical Research Council's Institute of Neurology in London, and her colleagues.

The researchers monitored the behaviour of mice infected with a prion protein to look for early indicators of the disease. Changes in the way the mice responded to their environment occurred before the onset of obvious signs of neurodegeneration, such as reduced grooming. The mice recovered their normal brain function if production of the naturally occurring protein that propagates the disease was switched off at this stage.

But the implications for treatment of the human prion disease, Creutzfeldt–Jakob disease, are uncertain. It's not clear whether human brains could recover, nor how levels of the prion protein could be lowered.

IMMUNOLOGY

The right kind of help

J. Exp. Med. doi:10.1084/jem.20061839 (2007)

So far they have only tested their hypothesis in mice, but researchers think they have identified a mechanism that could help to explain why men are less prone than women to developing certain types of autoimmune disease, such as mutliple sclerosis.

Lawrence Steinman of Stanford University Medical Center, California, and his colleagues studied a receptor known as peroxisome proliferator-activated receptor-α, which has been implicated in gender differences in lipid metabolism. The receptor is also expressed in the immune system's $CD4^+$ T cells.

The researchers showed that the receptor gene is sensitive to testosterone, and is expressed at higher levels in the T cells of male mice than in those of females.

$CD4^+$ T cells differentiate into different types of 'T-helper' cell. Expression of the receptor seems to direct differentiation away from the type that is associated with certain autoimmune diseases. Knocking out the gene in males made the symptoms of a mouse model of multiple sclerosis more severe.

PALAEONTOLOGY

Could the 'hobbit' hunt?

HOMO — J. Comp. Hum. Biol. doi:10.1016/j.jchb.2006.11.001 (2007)

Debate over the diminutive *Homo floresiensis* — believed to be a hobbit-sized species of hominid — has inspired a team at Washington University in St Louis, Missouri, to develop a method to estimate the size of hominid brain components from fossil skulls.

Researchers have questioned whether the small-brained *H. floresiensis*, which lived on an isolated Indonesian island until at least 12,000 years ago, would have been capable of creating tools, using fire and hunting, as some studies have suggested.

Glenn Conroy and Richard Smith looked at the volumes of 11 different brain components in 45 primate species to set limits on the size of each component as a fraction of overall brain size. The predicted bounds for the brain of *H. floresiensis* are not

consistent with the suggested behaviour, they conclude, unless the brain of *H. floresiensis* was functionally different from that of modern humans or chimpanzees. That possibility would force biologists to rethink how hominid brains evolved.

METAGENOMICS

Invisible communities

Science doi:10.1126/science.1133420 (2007)

Thousands of different microorganisms can share the same habitat, but these populations are hard to study because few individual species survive in cell culture. One way around this is to perform shotgun sequencing of the whole communitiy — an approach known as metagenomics.

Peer Bork from the European Molecular Biology Laboratory in Heidelberg, Germany, and his colleagues present a new bioinformatics approach to digging out biologically relevant details from such data.

They examined 31 'marker genes' — genes that are present in all microbes, but show variation between species — to probe the composition of four communities, including microbes from ocean water and from soil. This allowed them to show, for example, that some communities evolve faster than others and that some microbes have clear and persistent preferences for particular environments.

MATERIALS SCIENCE

Stiffer than diamond

Science **315,** 620–622 (2007)

For hardness and stiffness, it's long been thought that nothing beats diamond. But Roderic Lakes of the University of Wisconsin -Madison and his colleagues have made a material that is almost ten times stiffer, by embedding small particles of barium titanate in a matrix of tin.

Barium titanate can adopt different crystal structures. In the composite, it is trapped in a high-temperature form at temperatures below the usual transition point. This gives the inclusions the strange property of negative stiffness, meaning that they bend in the opposite direction to an applied force.

Warming the composite changes the balance between the mechanical properties of the inclusions and those of the matrix, which responds more conventionally to an applied force. At a particular temperature, these tendencies cancel each other almost exactly, and the material then scarcely deforms at all.

This extreme stiffness is remarkable, given that neither of the materials involved are especially stiff on their own. The composite isn't, however, expected to be particularly hard or strong.

ASTRONOMY

Super-Earth model

Astrophys. J. **656,** 545–551 (2007)

Recently discovered planets in solar systems beyond our own have masses just a few times that of Earth. What might these 'super-Earths' actually look like?

Diana Valencia and her colleagues at Harvard University in Cambridge, Massachusetts, do the maths. They assume that the planets are composed of concentric shells of similar compositions to Earth's layers, and solve equations for the density, gravity, mass and pressure in each as a function of radius.

Varying the proportion of core and mantle, or postulating a surface ocean, produces different solutions for the model planets' mean density and radius that could be compared with future observational data. For instance, if the radius of GJ 876d, the super-Earth reported in 2005, is more than 12,000 kilometres, then it almost certainly contains water, the authors say.

ZOOLOGY

Food poisoning

Proc. Natl Acad. Sci. USA doi:10.1073/pnas.0610785104 (2007)

They say you are what you eat. And that's especially true of the snake *Rhabdophis tigrinus* — zoologists have discovered that it eats poisonous toads and keeps their venom for itself.

R. tigrinus, which is found in Asia, eats a wide range of prey, including toads that secrete defensive poisons known as

D. A. HUTCHINSON

bufadienolides through their skin. When fed a diet featuring these toads, the snakes' venom glands fill up with a poison that seems to have been made from these chemicals, report researchers led by Deborah Hutchinson of Old Dominion University in Norfold, Virginia, and Jerrold Meinwald of Cornell University in Ithaca, New York.

The snakes store the venom in structures known as nuchal glands on the backs of their necks, underlying the ridge visible in the picture above, which they display if confronted by a predator. Snakes with no toads, and therefore no toxin, in their diets make these defensive displays less frequently.

JOURNAL CLUB

Axel Kleidon
Max-Planck Institute for Biogeochemistry, Jena, Germany

A biogeochemist finds inspiration for life on the ocean floor.

My research interests lie in understanding the interplay between the physical and chemical conditions that constrain life, and the feedback processes by which life shapes the Earth's environment.

I want to understand these interactions in terms of a thermodynamic hypothesis that states that systems dissipate as much energy as possible. Can life be seen as an emergent outcome of this tendency for the whole Earth system? To test this, one would need to show that it is possible to predict the emergence of life from the hypothesis, as well as its impact on Earth's early environment.

Two articles (M. J. Russell & A. J. Hall *GSA Memoir* **198,** 1–32; 2006, and M. J. Russell *Am. Sci.* **94,** 32–39; 2006) could provide a starting point. The authors give a detailed picture of the thermodynamics of life emerging at hydrothermal mounds on the ocean floor.

One of the earliest metabolic reactions would have involved the conversion of hydrogen, carbon dioxide and sulphur compounds into organic carbon, acetate and water. This would have happened in the hot, mineral-rich spring water seeping into the hollow mound.

But its influence would have been felt more widely. Removing sulphur from the environment would have changed atmospheric composition and cloud cover, affecting the amount of sunlight reaching the ground. And acetate may have served as fuel for methanogens, methane-producing organisms known to live in vents. Increased methane production would have raised its levels in the atmosphere, resulting in higher surface temperatures on Earth.

Quantifying these interactions should help us to understand whether the evolution of our planet emerged from general thermodynamic trends.

NEWS

Bush and Congress set debate on priorities

President George W. Bush's vision of America's science priorities, as revealed in his budget request for fiscal year 2008, is pretty much the same as it was last year. But one crucial thing has changed: the Democrats are now in charge of Congress, and they are the ones who get to decide whether his request becomes reality.

When it received the president's request on 5 February, Congress was already in the process of ladling out an extra $620 million for biomedical research in the current year, enough for the president's 2008 request to actually represent a cut. Bush has focused instead on increasing funding for the physical sciences under the second year of a 'competitiveness initiative' to stimulate investment in science and engineering. Overshadowing all such requests is the war in Iraq and the global war on terror, for which Bush is requesting $142 billion for 2008.

Meanwhile, US researchers continue to deal with the messy leftovers of last year's budget process, of which the biomedical science boost forms a part. The budget process ground to a halt in December when the old, Republican-led Congress left without finalizing budget bills for fiscal year 2007. Soon, though, the current Congress is expected to pass a sort of halfway budget for 2007, which would keep most agencies at their 2006 spending levels but make adjustments for some. The House of Representatives passed this bill on 31 January and the Senate is expected to do so by 15 February.

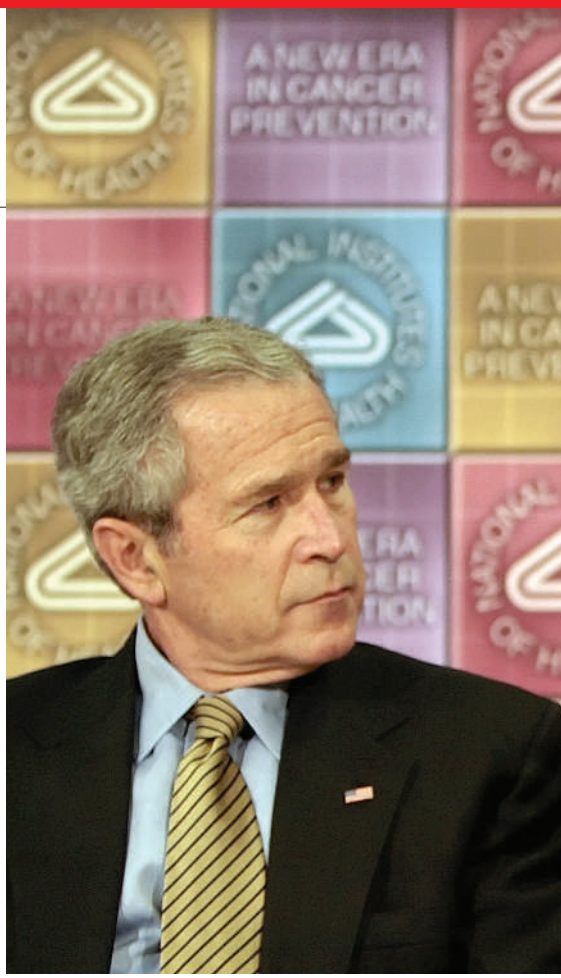
This week, Bush asked Congress to give the National Institutes of Health (NIH) $28.7 billion in 2008 — $232 million of it new money compared with 2006 levels. But of that, $200 million would have to be transferred to the Global Fund to Fight HIV/AIDS, Tuberculosis and Malaria. That means, in effect, that Bush's request would cut the agency budget by roughly half a billion dollars, or 1.7% in 2008.

Health check

Elias Zerhouni, the NIH director, said he was pleased with both the congressional action and Bush's budget request. "I personally feel that there has been a huge turnaround — that Congress has regained confidence in the NIH," he said. "The same thing is true of this administration."

Research advocates weren't so upbeat. David Moore, a lobbyist at the Association of American Medical Colleges, says that his group has been encouraged by Congress approving the extra money for 2007. "We're equally discouraged with the administration's budget," he adds, "which essentially puts the NIH right back in the hole again."

But Congress may boost the agency's budget further in 2008. The key congressional appropriators in both chambers of Congress are strong NIH advocates. And last month, Bush



NIH director Elias Zerhouni (right) feels President Bush's budget recognizes his agency's true role.

US SCIENCE AND TECHNOLOGY BUDGET (in millions of dollars)

Agency	2006 (actual)	2007 (likely)	2008 (president's request)	Notes
National Institutes of Health	28,359	28,979	28,700	An additional $158 million from other government sources, for diabetes research and the National Library of Medicine, will be provided in 2008 but is not included in the request at left.
National Science Foundation	5,590	5,916	6,430	Restored funds for 2007 may allow the United States to participate in the International Polar Year as planned. A new petascale computing centre is also in the works.
Department of Energy's Office of Science	3,633	3,796	4,398	Much of the proposed boost is related to the American Competitiveness Initiative to foster innovation. Also includes funds for several biofuel research centres.
NASA	16,273	16,247	17,310	Plans to replace the space shuttle are likely to have to move more slowly than planned because of expected cuts for 2007.
National Oceanic and Atmospheric Administration	3,851	3,851	3,812	The president called for more funding for ocean research in late January, but the agency would still face cuts under his 2008 request.
National Institute of Standards and Technology (core funding)	431	491	594	Although core funding, for science and related construction projects, is increased, the zeroing out of congressionally mandated projects or earmarks, means that the total 2008 budget goes down for the agency.
Environmental Protection Agency	7,619	7,619	7,200	The EPA's budget continues to be eroded.
US Geological Survey	958	978	975	$3 million of the increase for 2007 will be for ocean science programmes.
Department of Defense (basic & applied research)	6,405	6,893	5,785	The Defense Advanced Research Projects Agency budget for 2008 shrinks by 1% to $3.1 billion.
Homeland Security (science & technology)	1,467	848	799	Funds are shifting to a new office for detecting nuclear and radiological threats.

REWIRING BRINGS BACK TOUCH AFTER LIMB LOSS
Surgery opens door to prosthetics that can 'feel'.
www.nature.com/

REHAB. INST. CHICAGO

J. WATSON/AFP/GETTY

signed a law laying out broad directions for the NIH's future funding and administration (see *Nature* **444,** 796–797; 2006). The law recommends that Congress fund the agency at $32.8 billion in 2008. That might be pie in the sky, but it's a pie that Congress might at least cast a glance at in the coming months.

Getting physical

Meanwhile, the push is on to dramatically boost funding in the physical sciences. The president's request for 2008 includes $4.4 billion for the Department of Energy's Office of Science, which funds such fields as high-energy and nuclear physics. That would be a 16% increase over the tentative 2007 funding levels currently under consideration in Congress. The request includes $51 million for construction of the Linac Coherent Light Source at Stanford University in California, and $160 million to begin the US contribution to the International Thermonuclear Experimental Reactor (ITER), the $6-billion international fusion project that is slated to begin construction in Cadarache, France.

The competitiveness initiative would also provide a 21% jump in core research funding at the National Institute of Standards and Technology (NIST). That would bring NIST's total research budget to $501 million. The budget also calls for just under $94 million in funding for facilities that would go, in part, towards a new laboratory for atomic-scale electronics at the agency's campus in Boulder, Colorado. Overall, NIST's budget would drop 5% owing to the exclusion of several congressionally mandated projects, known as earmarks.

The competitiveness initiative also showers cash on the National Science Foundation (NSF), which would receive $6.43 billion — an increase of more than 15% over its 2006 funding. The money goes to particular directorates within the NSF, including engineering; computer and information science and engineering; and mathematics and physical sciences. This division worries Joel Widder, a lobbyist with the firm Lewis-Burke Associates in Washington DC and a former NSF staffer. "It is clear that the administration is deciding which directorates are part of the competitiveness initiative," he says. "I am concerned that this may become a more pronounced difference over time." He thinks the agency should have more control over how the money is divided up.

"In some ways, we are starting to pit sciences against one another, and I don't think that is helpful."

The competitiveness initiative has some Democratic friends on Capitol Hill, including Bart Gordon (Democrat, Tennessee), who chairs the House Committee on Science and Technology. And a similar programme to boost funding for the agencies was unveiled in 2005 by Nancy Pelosi (Democrat, California), who is now Speaker of the House.

But getting the substantial increases for the initiative through Congress may prove difficult. For 2007, congressional supporters were able to add only about half of what the president had requested for the initiative in its first year. That discrepancy may make it even harder to bolster the numbers further for 2008, says Michael Lubell, head of public affairs for the American Physical Society. "Clearly, it's a heavy lift," he says.

Space fight

Another agency that may see congressional wishes lined up against the president is NASA. Bush's request calls for $17.3 billion for the agency. That's a 6.5% increase over what NASA is likely to get if the 2007 spending bill passes the Senate, but the latest numbers don't offer much cause for optimism.

The spending bill that passed the House last week includes a $26-million cut for the agency from its 2006 levels, at a time when NASA is struggling to complete its space shuttle and space station programmes. "While the money being proposed for 2008 is back up to where it was planned to be, you can't make up for that year of inadequate funding," says John Logsdon, director of the Space Policy Institute at George Washington University in Washington DC.

NASA administrator Michael Griffin this week admitted that plans to develop the Orion capsule and Ares I rocket to carry astronauts into orbit may take longer than expected. "I am concerned about our ability to bring these new capabilities on-line by 2014," he said.

Within its science priorities, though, NASA has reshuffled funds in order to launch the Global Precipitation Measurement Earth-observing satellite in 2013, which is earlier than had been expected. And a servicing mission to the Hubble Space Telescope is set to go ahead in September 2008.

All the winners and losers in this year's budget race make Sam Rankin, director of the American Mathematical Society office in Washington DC, nervous. Take any agency or programme, he says, and "you do a chart of the differences from one year to the next and it looks like someone with bad teeth. In some ways, we are starting to pit sciences against one another, and I don't think that is helpful. We need something that is consistent and adequate." ■

Emma Marris, Geoff Brumfiel, Meredith Wadman and Lucy Odling-Smee

See Editorial, page 568.

www.whitehouse.gov/omb/budget/fy2008

Dig links Stonehenge to circle of life

A. STANFORD/AERIAL-CAM FOR NATIONAL GEOGRAPHIC

C. STEELE PERKINS/MAGNUM FOR NATIONAL GEOGRAPHIC

This settlement found near Stonehenge (inset) could have been home to the monument's builders.

When archaeologists announced last week that they had uncovered a major settlement at the site of Stonehenge, you could be forgiven for asking why it hadn't been found before. The legendary circle of standing stones, located in densely populated southern England, is a popular tourist attraction and one of the most famous prehistoric monuments in the world.

But on 30 January, archaeologists reported evidence that Stonehenge is just half of a bigger ceremonial complex. Excavations at a site called Durrington Walls, about three kilometres from Stonehenge, uncovered a cluster of houses and a vast stone avenue that connects the village to what would have been a circular arrangement of huge timber posts.

"There is a kind of complementarity between the timber circle and the stone circle of Stonehenge," says Julian Thomas of the University of Manchester, and a director of the Stonehenge Riverside Project, which excavated the site last year. Stone avenues connect both circles to the nearby river Avon, but Stonehenge is oriented to face midsummer sunrise and midwinter sunset, whereas the wooden circle would have faced midwinter sunrise and midsummer sunset.

Carbon dating suggests that the structures at Durrington Walls were built around the same time as Stonehenge, in 2600–2500 BC, supporting the idea that the villagers helped to build the stone monument. As well as being the "richest site archaeologically speaking, it's also the filthiest we have ever seen", says project director Mike Parker Pearson of the University of Sheffield, UK. Some 50,000 animal bones found at the site suggest that people went there for "feast and fun". He speculates that the stone and wooden monuments played contrasting roles in cult practices — the wooden one being used to celebrate life, and the stone one to celebrate death.

"There is a kind of complementarity between the timber circle and the stone circle of Stonehenge."

How did archaeologists miss all this material before? "Because it's a World Heritage Site, there has been a bit of a disincentive for people to go and excavate there," explains Timothy Darvill, an archaeologist at Bournemouth University, UK, who was not involved in the latest dig. "In the past, the thrust has been to conserve and protect such areas."

But in recent years, attitudes have been shifting, says Darvill. "People have realized that if you do nothing, you find nothing." That shift in philosophy seems to be paying dividends. ■

Lucy Odling-Smee

For a more detailed story and an animation of the site, see http://tinyurl.com/2wytlr.

Virus paper reignites prion spat

Mention the name Laura Manuelidis to a prion researcher, and you will almost certainly get a strong reaction. For years the Yale University neuropathologist has ruffled feathers by arguing that a virus causes neurological conditions such as Creutzfeldt–Jakob disease in humans and scrapie in animals, rather than an infectious protein, or 'prion', as most researchers believe.

That argument is decades old, and the prion hypothesis has been steadily winning converts. But the publication of a paper from Manuelidis in the *Proceedings of the National Academy of Sciences* last week served as a reminder that the debate is not over[1]. The paper, which shows pictures of virus-like particles in infected cells in culture, hardly overturns the prion hypothesis, but it does show how hard it can be to pin down the source of a disease.

There was a time when the tables were reversed. Stanley Prusiner from the University of California, San Francisco, fought for years to win acceptance for his idea that a misshapen protein was the cause. When he got a Nobel prize in 1997, the fight for public opinion was won but scientific skirmishes continued.

Researchers had shown that infecting animals with purified prions could cause disease, but some still questioned whether a virus particle or two could have slipped undetected through the purification process. But in 2004 Prusiner's lab showed that prions synthesized in a test tube — and so free of virus — could cause disease in mice[2].

A year later, Claudio Soto, a neurologist at the University of Texas in Galveston, selectively amplified prions from diseased mouse brains. He diluted his solutions until they could not contain any molecule from the brain sample other than prions. The solutions also infected mice[3].

That work silenced much of the debate, says Neil Cashman, scientific director of PrioNet Canada in Vancouver. "I really feel that these viral people are swimming upstream," he says. "But some highly regarded people do not think the prion hypothesis has been proven."

Prusiner's work was important but flawed, says Robert Rohwer, a neurovirologist at the Veterans Affairs Medical Center in Baltimore, Maryland. The amount of synthetic prion needed to infect the mice was a million times higher than would be required using brain isolates. And the mice were engineered to produce high amounts of the prion's natural form, making them prone to developing prion disease spontaneously.

> "It's not very politically correct if you don't believe in prions."

If the synthetic prion can be shown to infect normal mice, Rohwer says he will be convinced. "If it's successful and can be reproduced by others, then it really does close the issue as far as I'm concerned," he says.

Prusiner's results have not yet been verified by other labs, and some researchers have complained that the transgenic mice he used were not made available. Ilia Baskakov, a lead author on Prusiner's paper who is now at the University of Maryland in Baltimore, says they used transgenic mice because these had the strongest response, but he hopes to replicate the work in wild-type animals soon.

Baskakov says the synthetic prion probably had low activity because it was not folded into an optimally infectious form. "This is an extremely complex problem," he says. "I would not expect that somebody would just generate 100% infectious prions in their first experiment."

Meanwhile, Bruce Chesebro, chief of the Laboratory of Persistent Viral Diseases in Montana, is concerned that the infectivity of Soto's prion solutions didn't increase proportionally with their concentration. Soto says this was unique to the prion strain used for those experiments, and that subsequent, unpublished, work with other prions has shown that virulence does increase with concentration.

For some researchers, these are mere details, and they clearly wish that Manuelidis and her arguments would just go away. Two refused to speak to *Nature*, because they did not want to play any role in publicizing her work. Charles Weissmann of Scripps Florida in Jupiter says the evidence in favour of prions is incontrovertible. Although Manuelidis claims in her paper that virus-like particles from infected cells can infect other cells in culture, Weissmann says the findings are meaningless until she can purify her particles and show that they can cause disease and that they are free of the prion protein — something he does not expect to happen.

Manuelidis believes that her criticism of prions has made it harder for her to publish in top journals. "It's not very politically correct if you don't believe in prions," she says. "It means that you won't get any grant money."

Although many disagree with her, several prion researchers say that Manuelidis's scepticism keeps the field on its toes. "To me the prion issue is settled, but it's good to have this type of person," says Soto. "This issue is so important that you have to be very rigorous." ■

Heidi Ledford

1. Manuelidis, L., Yu, Z.-X., Banquero, N. & Mullins, B. *Proc. Natl Acad. Sci. USA* doi:10.1073/pnas.0610999104 (2007).
2. Legname, G. *et al. Science* **305,** 673–676 (2004).
3. Castilla, J., Saá, P., Hetz, C. & Soto, C. *Cell* **121,** 195–206 (2005).

NUMBER CRUNCH

98 billion is the mass, in tonnes, of one square kilometre of the United Kingdom, say geologists who have weighed up Europe's countries by evaluating the thickness of Earth's crust in different regions.

112 billion is the mass of the same area of Austria, which, thanks largely to the Alps, is Europe's densest country.

78.4 billion is the corresponding figure for Germany, which, despite being an economic heavyweight, is less well endowed in the geological stakes.

SCORECARD

▲ **Eagle-eyed diagnosis** An unnamed Irish surgeon has been congratulated for spotting a tumour in government minister Conor Lenihan's cheek while watching him on television. He called Lenihan's office to urge him to see a specialist; the tumour has now been successfully removed.

▼ **Negative calories** Coca-Cola and Nestlé are being sued by advocacy group the Center for Science in the Public Interest over claims that Enviga, the 'negative-calorie' drink, can help people lose weight by speeding up their metabolism.

COCA-COLA

WORDWATCH

Google bombing
The trick used by webmasters to code links into their websites in sufficient numbers to manipulate search engines, ensuring, for example, that the term 'miserable failure' links to George W. Bush's biography. Google has pledged to tackle the practice, leading to accusations that it is censoring its results.

ON THE RECORD

"All food for purchase in the UK is 110% safe."

Nigel Horrox, president of the British Veterinary Poultry Association, seems to think that the arrival of the H5N1 flu virus at a British turkey farm has boosted food quality to an unprecedented degree.

Sources: Ananova Science, The Guardian, CSPI, UPI.com, Science Media Centre

SIDELINES

Africa pursues goal of scientific unity

ADDIS ABABA

African heads of state and ministers ended their summit in Addis Ababa, Ethiopia, on 30 January amid a blaze of initiatives on education, science and sustainable development — once they had stopped talking about football.

Decisions made at the three-day summit of the 53-member African Union (AU) include plans for three new funding agencies, an Africa-wide system of intellectual-property protection, and a 20-year strategy for biotechnology. Africa's leaders also asked the AU secretariat to produce an assessment of alternative energy sources in the continent, particularly biofuels, before their next meeting in July.

The meeting was driven by the AU's science team, headed by Nagia Essayed, and the head of her secretariat, Botlhale Tema. By lunchtime on the first day, both women were looking dejected. The themes of science and the environment had been overshadowed by the glitzy launch that morning of the International Year of African Football. Television coverage of the summit, moreover, focused on political conflicts between member states and disagreements over the election of the AU's chair.

During lunch, Essayed sat with several science ministers and their advisers to discuss how to steer the presidents back to the summit's main business. Calestous Juma, co-chair of the AU's biotechnology advisory panel, had a lighthearted suggestion to link science and football: "Countries that export soccer balls need to have expertise in precision engineering."

But the breakthrough happened the following morning when Nicholas Stern, economic adviser to the UK government and author of an influential report on climate change, gave a speech on the causes and effects of climate change. This provoked an angry response from several presidents, including Yoweri Museveni of Uganda and Blaise Compaoré of Burkina Faso, who saw it as a hypocritical lecture on how to be clean from a representative of one of the world's major polluting countries. Museveni said he regarded global warming as an act of aggression by developed countries against Africa.

But Stern's stark message made clear how much African countries stand to suffer from global warming. Essayed says that afterwards it was easier to convince heads of state that Africa urgently needs stronger scientific capacity — not least to help the continent adapt to rising temperatures. "I felt that we got more support for science after what happened during the climate-change debate," she says.

K. PRINSLOO/AP

UN secretary-general Ban Ki-moon addresses the 8th African Union summit in Addis Ababa.

The most ambitious commitment came from Paul Kagame, president of Rwanda. In his keynote speech, Kagame pledged to double his country's spending on research and development to 3% of its national income — the average for developed nations — within five years.

The heads of state gave the go-ahead to the Green Wall for the Sahara, a ten-year plan to green the Sahara/Sahel region. And they agreed that work should begin on a 20-year biotechnology strategy. They hope to reconcile scientists, environmentalists, industry and church groups, who have been arguing over genetically modified crops, and find a common approach. The decision may disappoint Germany's scientific and technical aid agency GTZ, which is funding a separate Africa-wide project to develop some tough regulations on biotechnology.

> **"I felt we got more support for science after what happened during the climate-change debate."**

The three funding agencies to be set up are an education fund, an environment fund based on the successful United Nations Global Environment Facility, and a petroleum fund that will enable oil-rich countries to provide a financial cushion for poorer nations when oil prices are high.

The reforms to intellectual property protection are centred on creating a pan-African organization to replace the two existing organizations that do this job — one for the 16 English-speaking countries, and one for the 16 French-speaking ones. The new body, vigorously lobbied for by the Geneva-based World Intellectual Property Organization, will also provide Africa-wide intellectual-property protection for north Africa's Arabic-speaking countries. The AU secretariat has until July to come up with a progress report.

Initiatives that failed to get the nod include a council of presidents to monitor the implementation of the AU's scientific commitments. This was rejected on the advice of science ministers who felt that this was their job, for which they have recently set up the African Ministerial Council on Science and Technology.

A second initiative that failed, for now, was an Africa-wide science and innovation fund. This fund may yet happen, according to Essayed. Heads of state asked for modifications, and the plan will be presented again at a future AU summit. "I am going to nag them until we get this one through," she says.

Beyond science and technology policy, the heads of state also moved ahead with plans for an African parliament, a central bank, a monetary fund and a court of human rights. These are in line with the AU's desire for complete political and economic integration, leading to what it calls a United States of Africa. ■

Ehsan Masood

Trials halt after gel found to increase HIV risk

Microbicidal gels are one of the great hopes for reducing rates of HIV in women as they are cheap, easy to apply and don't require the consent of male partners. But the news that one of the gels being tested actually increases the risk of infection has shocked scientists and halted clinical trials.

The cellulose sulphate gel, called Ushercell and made by Polydex, based in Toronto, Canada, was in phase III clinical trials involving more than 1,330 women in South Africa, Benin, Uganda and India. The trials were stopped when researchers realized that the infection rate was higher in women using the gel, although the exact figures have not yet been released. As laboratory tests showed that the gel is active against HIV, scientists are at a loss to explain the result. Principal investigator Lut Van Damme calls the findings "unexpected and disappointing".

A separate phase III trial on the same gel in Nigeria has also been stopped as a precautionary measure. Cellulose sulphate is one of four potential anti-HIV microbicides undergoing phase III trials.

Cloners take different views over best recipe

Overturning conventional wisdom is never simple, and a recent claim that it is easier to make cloned animals from cells that have already become specialized than from adult stem cells has provoked a stern riposte.

Xiangzhong Yang and his team at the University of Connecticut, Storrs, reported the finding last October (L.-Y. Sung *et al. Nature Genet.* 38, 1323–1328; 2006). It was surprising because fully differentiated cells are generally thought to be unable to produce other cell types, whereas stem cells are more flexible. Two stem-cell experts have now published a letter of dissent. "One wonders whether the evidence presented in this paper justifies the sweeping conclusion," write Rudolf Jaenisch of the Whitehead Institute, Cambridge, and Konrad Hochedlinger of the Harvard Stem Cell Institute in Boston (*Nature Genet.* 39, 136–137; 2007).

Jaenisch and Hochedlinger take issue with the way Yang's group measured cloning efficiency, and question whether the cells used were in fact terminally differentiated. They also challenge Yang's claim that he is the first to show that clones can be made from terminally differentiated cells. Yang defends his methods and conclusions, and says further experiments — now under way — will prove him right.

Campaigning for open access gets easier

Organizers of an Internet petition supporting free access to scientific results have hit a nerve. Nearly 18,000 people and institutions have signed up since 17 January. "I think it is one of those interesting viral network sort of effects," says David Prosser, director of the Scholarly Publishing and Academic Resources Coalition, Europe (SPARC). "It is everything from graduate students to Nobel prizewinners."

The petition supports the recommendations of a European Union study of January 2006 which called for "public access to publicly funded research results shortly after publication", and urges the European Commission (EC) to make access mandatory for research that it funds. The group intends to present the petition to the EC at a 15 February conference in Brussels on scientific publishing.

Japanese nets threaten grey whale's survival

D. WELLER

Japan is having too many accidents involving the critically endangered western grey whale (*Eschrichtius robustus*), the World Conservation Union (IUCN) warned on 1 February.

A young female grey whale found dead in a fixed net off the northeastern coast last month was the fourth since 2005. Only some 120 western grey whales (pictured) are thought to be left, of which 20–25 are reproductive females. The current death rate is driving the species to imminent extinction, says Carl Gustaf Lundin, head of the IUCN's Global Marine Programme, based in Gland, Switzerland. "It is essential that the causes of net entrapments are investigated thoroughly so that remedial action can be taken," he says. Japan's Fisheries Agency will order more detailed analyses of any further deaths, but says it has no plans to take action to prevent entrapments.

LAWRENCE BERKELEY NATL LAB.

Glory days: physicists Edwin McMillan and Edward Lofgren at the Bevatron in 1963.

Demolition looms for 'landmark' Bevatron

The Bevatron, the renowned particle accelerator in Berkeley, California, has had its landmark status confirmed — but can be knocked down anyway.

On 1 February, the city council upheld its earlier ruling allowing Lawrence Berkeley National Laboratory to demolish the accelerator, where the antiproton was discovered in 1955. The lab wants to tear down the ageing facility to make way for other projects, but local activists, who opposed demolition on environmental grounds, argued that it should be left as a landmark (see *Nature* 442, 612; 2006).

The council did, however, recommend that the lab erect a memorial to mark the achievements of the machine, which operated for more than 40 years. Demolition is not expected before 2008.

Academic starts hunger strike for tenure at MIT

Biologist James Sherley, a black associate professor at Massachusetts Institute of Technology (MIT) in Cambridge, began a hunger strike on 5 February in protest at being denied tenure because of what he claims is racial bias at the institute.

MIT stands by its decision to deny Sherley tenure, and insists that the decision process was fair. Provost Rafael Reif said in an e-mail to students and faculty members that he and MIT president Susan Hockfield "are deeply committed to removing barriers that may exist for under-represented minority faculty members", and that a committee will be formed to study potential racial bias in the institute's hiring and career-advancement procedures.

But Sherley, a stem-cell researcher in the university's biological engineering department, is unimpressed. "I don't think the committee exists," he says. Sherley has begun his hunger strike in the corridor outside the provost's office.

SPECIAL REPORT

From words to action

The scientific case for global warming is overwhelming. So what next for the IPCC? Helping policymakers decide what to do now may require radical reform, reports **Jim Giles**.

The disturbing predictions about global warming in the latest report from the Intergovernmental Panel on Climate Change (IPCC) mark a turning point. That's not because of the figures themselves, which are largely in line with previous IPCC forecasts, but because the science behind them is now certain enough to make a serious response from policymakers almost inevitable. The debate is no longer about whether we can believe the numbers, but what we should do about them.

And so the report, released in Paris on 2 February (see 'Behind the scenes'), may be the harbinger of another change. As the debate evolves, say climate researchers, so must the IPCC. Many now feel that its weighty structure — the latest report has more than 1,200 authors and reviewers — is no longer useful. If the panel is to guide policymakers in the future, it must slim down, and become more focused on producing data that politicians can use. "The IPCC needs a complete overhaul," says Mike Hulme, director of the Tyndall Centre for Climate Change Research in Norwich, UK. "The structure and process are past their sell-by dates."

This might seem contrary, given the panel's enormous influence. But the proposed shift is partly a recognition of the success that climate scientists and the IPCC have had in revealing the scale of the problem. The latest report, the panel's fourth, is peppered with predictions classified as "highly likely" and "unequivocal". It notes with over 90% certainty that recent temperature increases are driven by human activity, and describes detailed impacts in more places than ever before, from melting ice sheets to shifting wind patterns.

Some of that confidence comes from the accuracy of previous IPCC predictions, such as estimates made from 1990 onwards, that global temperatures would rise by between 0.15 °C and 0.3 °C per decade. Temperatures have climbed steadily since: the ten hottest years on record all postdate 1990, and the rate of warming, 0.2 °C per decade, fits the initial prediction.

More sophisticated models and mounting observational data have also reinforced simulations of future climate. For example, the new report is the first to go into detail on how warming will affect the carbon cycle. The previous report, released in 2001, was unable to say whether rising temperatures would further increase atmospheric carbon levels by speeding up the decay of organic matter in soils, or cut levels by promoting plant growth. It's now clear that the former effect dominates, says Peter Cox, a climate modeller and IPCC author based at the University of Exeter, UK. "All the models give positive feedback."

"The IPCC needs a complete overhaul. The structure and process are past their sell-by dates."

The report is the first to give a best estimate — 3 °C — of 'climate sensitivity': the global mean temperature rise resulting from a doubling of carbon dioxide levels. The range of possible values has also been tightened, to 2.0–4.5 °C. This feeds into predictions of the temperature increase over this century, which is now given as 1.1–6.4 °C, compared with 1.4–5.8 °C last time. Where we end up on that scale will depend mostly on how much fossil fuel the world burns.

"The report is an unequivocal set of evidence on how we are affecting our planet," concludes Achim Steiner, executive director of the United Nations Environment Programme. "The focus of attention must now shift to what we are going to do about it."

In the short term, attention will focus on further IPCC publications. The working group behind the current report looked only at the physical basis of climate change, but the IPCC has two other working groups that focus on the impact of climate change and what can be done about it. These will issue their 2007 conclusions in April and May, respectively.

After that, the decision about what to do rests with politicians. And if climate scientists are to inform those decisions they may need to reform the IPCC. Publishing coordinated and lengthy assessments at six-yearly intervals makes it difficult for authors to deliver timely advice and deters those from different working groups from collaborating. That's a problem, says Jonathan Overpeck, an IPCC author and palaeoclimatologist at the University of Arizona, Tucson, because producing research that can inform policy requires social and natural scientists to work together.

For example, the new report contains detailed regional predictions, such as that the western United States will warm on average by 5 °C by 2100. Such predictions will be needed regularly in the future, says Overpeck, and must include details such as the impacts on agriculture and water resources. Hulme agrees,

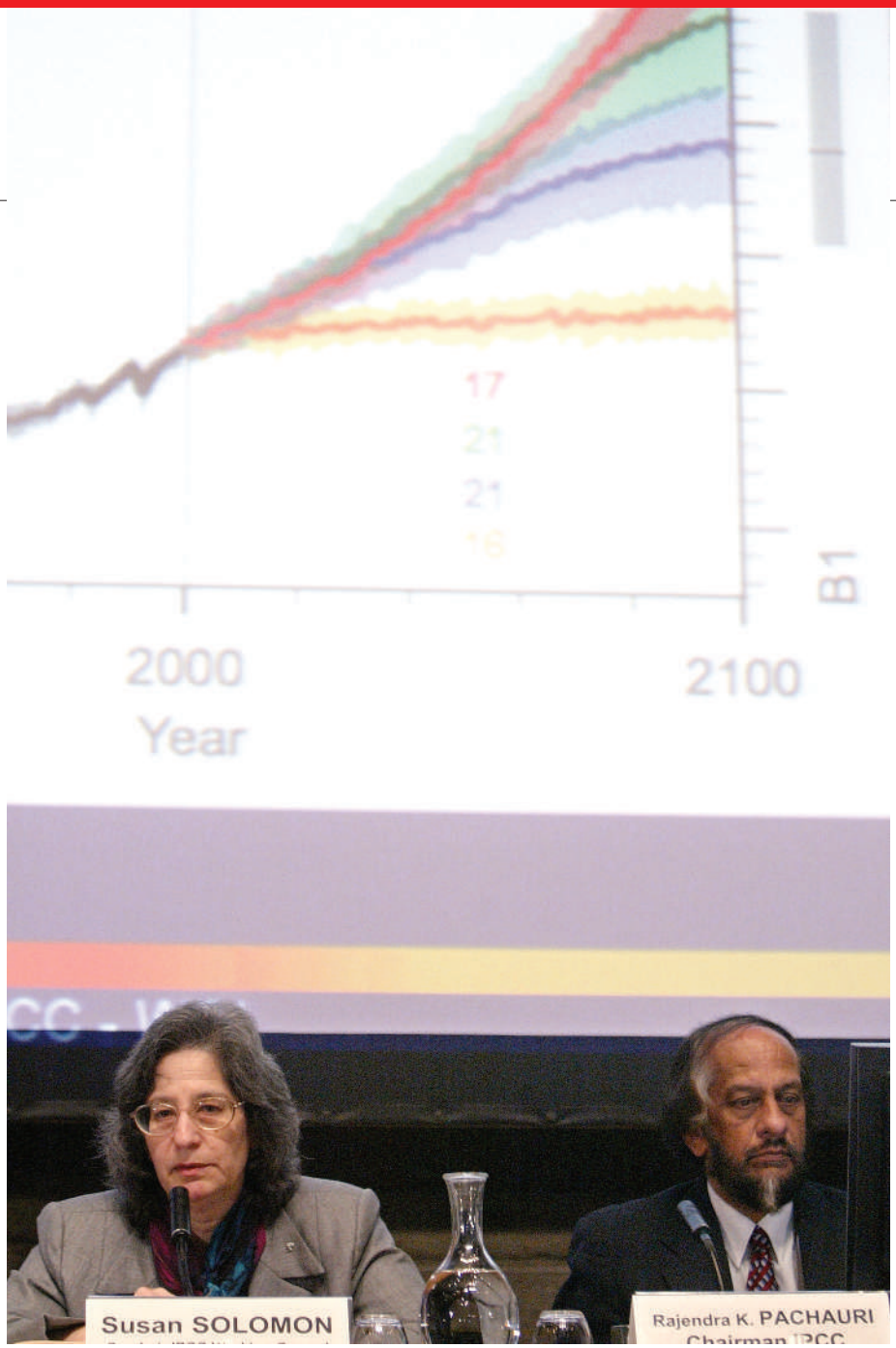
Behind the scenes

With the world's media ready to leap on the report's long-awaited conclusions, deciding the final wording of its summary was a sensitive process. After climate scientists completed a draft document last year (see *Nature* **441,** 6–7; 2006), political representatives spent the Paris meeting negotiating the summary line by line, with researchers in attendance to make sure vested interests didn't distort the science.

The consensus from the scientists is that the process worked well, with the final summary a fair representation of their conclusions. The United States has in the past been accused of seeking to play down some of the more alarming conclusions, but no such charges were levelled on this occasion.

There were a couple of sticking points. The Chinese and Saudi Arabian delegates annoyed scientists by insisting that a line stating that man-made warming "is at least five times greater than that due to solar output changes" was cut. But the data remain in the report.

Another hot topic was the contribution that melting polar ice makes to sea level rise. The publication of new papers on the subject, one on the day before the report itself (see page 580), prolonged the debate. But some researchers still felt the new results weren't properly considered, and that the final figures, which predict a sea-level rise of 0.18–0.59 metres by 2100, are too conservative. **J.G.**



E. CLEMENT/UPI PHOTO/NEWSCOM

Bearing bad news: senior IPCC figures Susan Solomon and Rajendra Pachauri unveil the panel's report.

adding that as the emphasis shifts from talk to action, IPCC reports will need to come in regular "digestible chunks" for policymakers.

Climate modellers must also aim to produce information that can inform policy directly, says Kevin Trenberth, a climate researcher at the National Center for Atmospheric Research in Boulder, Colorado. That means filling the gap between long-term simulations looking decades ahead and weather forecasts, to provide predictions for the next 20 or so years. "Some of what is in the research domain needs to become operational," argues Trenberth.

Rajendra Pachauri, chair of the IPCC, says he will circulate a document on possible reforms in the next couple of months, adding that he supports the suggestions made by the report's authors. Any change needs the approval of the political representatives who sit on the IPCC plenary and meet once or twice a year.

But some urge caution. The IPCC already produces reports from across the working groups, such as its 2005 study on carbon capture and storage, points out Halldor Thorgeirsson, a deputy executive secretary at the United Nations Framework Convention on Climate Change, the body that oversees the Kyoto Protocol. And the IPCC process drives research — because climate modellers pool resources when a report is coming up. That informs all three working groups. "We can't allow the scientific activities that underpin the assessments to dry up," says Thorgeirsson. ■

See Editorial, page 567

WHAT THEY'RE SAYING

"This may be remembered as the day the question mark was removed from whether human activity has anything to do with climate change."
Achim Steiner, head of the United Nations Environment Programme

"Now is not the time for half measures. It is the time for a revolution."
French president Jacques Chirac

"The question is, what can we do now? There's very little we can do about arresting the process."
Anote Tong, president of the Pacific island nation of Kiribati

"This should compel all of us towards action rather than the paralysis of fear."
Martin Rees, president of the United Kingdom's Royal Society

"Now it's time for us — the policymakers — to do our jobs."
Bart Gordon, Democratic Congressman from Tennessee and chair of the US House Committee on Science

"This is a group of climate experts attempting to reach a scientific consensus. It doesn't commit governments to any course of action."
Pradipto Ghosh, senior official at India's Ministry of Environment and Forests

"For sure, humans cause global warming!"
Headline from China's Xinhua news agency

"Let's be realistic. You can only run power stations in a modern Western economy on fossil fuel, or, in time, nuclear power."
Australian prime minister John Howard, whose country has not ratified the Kyoto protocol

"Those who continue to ignore the threat will be doing the greatest disservice imaginable to current and future generations."
Marthinus van Schalkwyk, environmental affairs minister for South Africa

Unusual local climatic events, such as the freezing temperatures last month in California, are hard to predict by global climate models.

G. KAZANJIAN/AP

What we don't know about climate change

The 4th Assessment Report from the Intergovernmental Panel on Climate Change (IPCC) has a finely calibrated lexicon of certainty. "Virtually certain", it blares when it assigns a 99% probability to hot days getting hotter and more frequent. "Very likely", or more than 90% probable, are heavier rains. And so on down the list — including the wishy-washy "more likely than not" when assigning a greater than 50% probability, such as the chance that human activities are affecting the intensity of hurricanes.

Such care is crucial in a field that is still, in some areas, shot through with uncertainty. The IPCC has gone far in tightening up some key scientific unknowns about climate change (see page 578), but many still remain. Some conclusions — such as the effect on particular regions of the world, or exactly how much sea level will rise — remain more uncertain than others. This means that there's plenty of work left for the climate scientists on whom the IPCC process depends.

Perhaps most critically, researchers know relatively little about feedback effects that might enhance — or weaken — the pace and effects of climate change. The complex flow of carbon between soils, plants, the oceans and the atmosphere is still being pinned down by large-scale climate experiments. Some experts predict that, in a warmer world, ecosystems that are currently sinks for carbon, such as the Arctic tundra, may turn into carbon sources. But no one can yet accurately predict how this might pan out, and feedbacks among land and air could end up putting far more carbon dioxide into the atmosphere than currently forecast.

"The IPCC report is not the leading — or bleeding — edge of the science."

Other big unknowns are the effects of the take-up of carbon dioxide by the oceans, which removes the gas from the atmosphere and locks it away in the calcium carbonate of the shells and skeletons of marine organisms. Higher levels of atmospheric carbon dioxide are expected to make the seas more acidic and slow down the rate of calcification, ultimately reducing the ocean's ability to absorb more carbon dioxide. But precisely how the biology of marine creatures would play into that effect is unknown. Nor is it known how changes in plankton composition and coral reefs, for example, might affect carbon dioxide concentrations.

Pinning down biological feedbacks will be critical for future reports, says Richard Bellerby, a chemical oceanographer at the Bjerknes Center for Climate Research in Bergen, Norway. "We're going blindly into the future," he says.

Another major source of uncertainty — and of debate at the Paris meeting where the IPCC report was finalized — is the rise in sea level. In 2001, the IPCC predicted a rise of between 9 and 88 millimetres by 2100, as a result of melting ice caps and the thermal expansion of the ocean. This time around, the group has narrowed that range to between 19 and 58 centimetres. But some scientists say that this is an underestimate.

Stefan Rahmstorf, an oceanographer at the Potsdam Institute of Climate Impact Research in Germany, believes, for instance, that global sea level could rise by much more than that. In a paper published online the day

CLIMATE CHANGE
Find all our news on global warming
http://www.nature.com/news/infocus/climatechange.html

before the IPCC report's release, Rahmstorf and his colleagues argue that sea-level rises will be close to the worst-case predictions of climate models (S. Rahmstorf *et al. Science* doi:10.1126/science.1136843; 2007). "If anything, the IPCC has been conservative," he says.

Key sticking points include the inability of global climate models to produce the amount of sea-level rise observed over the past couple of decades and whether ice flow at the bases of glaciers is accelerating or not. How volatile Antarctic and Greenland glaciers might become in a warmer world is therefore pretty much guesswork.

For the first time, the IPCC report predicts how changing climate might affect particular regions of the world. But these forecasts are only in their infancy, modellers warn. For some areas, models predict specific and well understood effects, such as hotter summers in Spain and smaller snowpacks (the accumulation of snow each season) in the Rocky Mountains in the United States. But improved analyses that incorporate clouds, snow and ice into the models must be developed if regional predictions are to become more accurate, says Rasmus Benestad, a climate modeller at the Norwegian Meteorological Institute in Oslo.

Extreme weather is another example of the remaining uncertainties. Climate researchers believe that storms and heavy rainfall will become more frequent as the planet warms. But pinning down where and when that might happen is not so simple.

In the tropics, rising sea-surface temperatures can be linked in a relatively straightforward manner to storm formation, and the case for more intense storms seems more or less settled. But in the mid-latitudes, where atmospheric processes are more complex, some climate models predict more storms whereas others do not.

Improving the models, experts say, requires better data. Gaps and errors in observations are attributable to many causes: snowfall gauges that ice up, oceanographic floats that get lost, and changeovers in satellites that throw off carefully calibrated trends, to name but a few. Cloud and storm records urgently need to be reprocessed using uniform techniques, says Kevin Trenberth, a senior scientist at the National Center for Atmospheric Research in Boulder, Colorado, and coordinating lead author of the report's chapter on surface and atmospheric change.

"The IPCC report is a consensus report, and one that develops over nearly three years," he says. "This means that it is not the leading — or bleeding — edge of the science." ■

Quirin Schiermeier

Data keep flooding in

The 2007 report of the Intergovernmental Panel on Climate Change (IPCC) represents the work of thousands of researchers, compiled and summarized by hundreds of climatologists. Nominally, the cut-off for inclusion in the assessment was the end of 2005, allowing a year for the panel to make sense of the vast tracts of data. But notable research arising after that date will not have escaped attention. Here is a round-up of some of the most prominent studies.

Greenland ice
Greenland is losing ice at an ever-increasing rate, according to data from the GRACE gravity-measuring satellite (J. L. Chen, C. R. Wilson & B. D. Tapley *Science* **313,** 1958–1960; 2006).

Antarctic air
Weather balloons reveal that the troposphere above Antarctica has warmed by 0.5–0.7 °C per decade over the past 30 years, although it is not clear why (J. Turner, T. A. Lachlan-Cope, S. Colwell, G. J. Marshall & W. M. Connolley *Science* **311,** 1914–1917; 2006).

Established forests
Old forests keep soaking up atmospheric carbon long after they reach maturity, according to measurements from China. Soil carbon in a forest reserve in Guangdong increased by 68% in 25 years (G. Zhou *et al. Science* **314,** 1417; 2006).

Atlantic currents
The Gulf Stream, which brings heat from the tropics to the North Atlantic, weakened by 10% between 1200 and 1850, during the cold spell known as the Little Ice Age. The authors suggest that this demonstrates the link between this ocean current and temperatures in northern Europe (D. C. Lund, J. Lynch-Stieglitz & W. B. Curry *Nature* **444,** 601–604; 2006).

Atlantic hurricanes
Rising sea-surface temperatures correlate strongly with the observed increase in the number of category 4 and 5 Atlantic hurricanes between 1970 and 2004. Other factors that affect hurricane formation, such as wind shear, do not seem to have increased in line with the upward trend (C. D. Hoyos, P. A. Agudelo, P. J. Webster & J. A. Curry *Science* **312,** 94–97; 2006).

River runoff
More carbon dioxide in the atmosphere leads to plants losing less water by transpiration, a model suggests. This could affect the amount of fresh water available for human use. (N. Gedney *et al. Nature* **439,** 835–838; 2006).

Polar temperatures
Bubbles dating back 150,000 years in an Antarctic ice core show that warming events have tended to seesaw back and forth between the poles (EPICA Community Members *Nature* **444,** 195–198; 2006).

Ocean temperatures
The upper layers of the oceans cooled, on average, between 2003 and 2005. Factoring in this downturn, the rate of warming in these layers between 1993 and 2005 was equivalent to 0.33 watts per square metre over the whole of the planet's surface (J. M. Lyman, J. K. Willis & G. C. Johnson *Geophys. Res. Lett.* **33,** L18604; 2006).

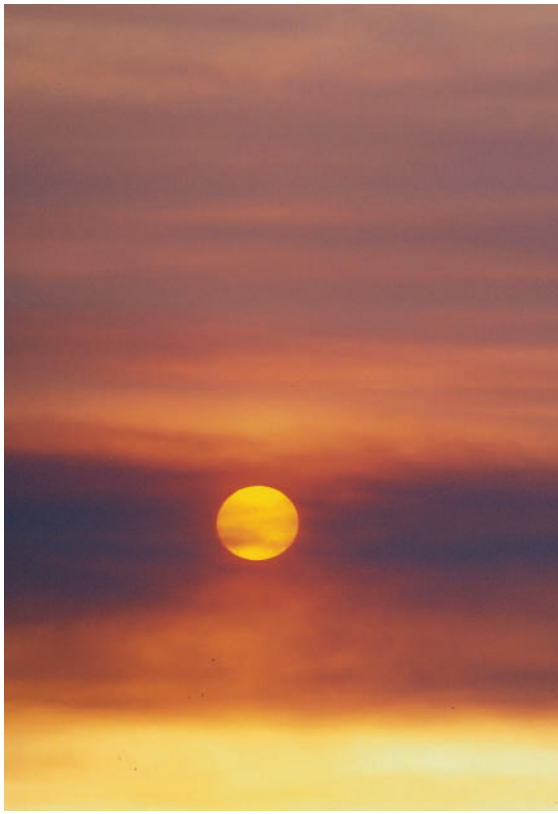
Sea levels
If the rate of sea-level rise is proportional to the global rise in temperature since pre-industrial times, sea levels could rise by up to 1.4 metres by 2100 (S. Rahmstorf *Science* **315,** 368–370; 2007).

Michael Hopkin

N. COBBING/STILL PICTURES

Greenland ice is melting faster than before.

Climate sceptics switch focus to economics



Reports by the Intergovernmental Panel on Climate Change (IPCC) are held out as a model of consensus science, with thousands of international scientists coming together to present the most detailed look ever at a single scientific topic. Yet a consensus among most of the world's researchers does not mean that everybody agrees.

"I am one of the 2,000 with their names on [the assessment], but don't sign me up for that catastrophic view of climate change," says John Christy, a climatologist at the University of Alabama in Huntsville and a contributing author to the report.

"Their argument continues to shift. That makes it clear that the issue is not the science."

And outside the IPCC process there remains a dwindling band of climate sceptics, those who argue that global warming is not linked to human activity and that it would be rash to take drastic action to cut carbon emissions. The focus of these arguments, however, has shifted noticeably since the previous IPCC report was published six years ago. Many of the scientific uncertainties the sceptics have seized on are no more.

"Their argument continues to shift," says Naomi Oreskes, a geologist and science historian at the University of California, San Diego. "That makes it clear that the issue for them is not the science. Whatever the science is, they will try to find ways to question it."

The previous IPCC report, for instance, gained infamy for featuring in its summary for policymakers the 'hockey stick' palaeotemperature graph. This shows a sharp rise in temperatures at the end of the last millennium that forms the blade of the hockey stick. Sceptics, notably economist Ross McKitrick of the University of Guelph in Canada and minerals consultant Steven McIntyre, have spent years working to discredit the statistical analysis and temperature proxies that were used to create the graph.

But a host of other studies, including a 2006 review by the US National Academy of Sciences, has reaffirmed that the past decade has seen an unprecedented rise in global temperatures. "This is just one of many lines of evidence," Michael Mann of Pennsylvania State University in University Park told an audience of congressional aides in Washington DC last week. Mann is the originator of one version of the graph.

Another key argument for sceptics has been the apparent discrepancy between warming at Earth's surface and temperatures in Earth's lower atmosphere, which satellite records suggested have cooled over the past several decades. But in 2005, a major US report, commissioned specifically to look into this question, concluded that the original data had been

What price a cooler future?

Last year's review on climate change by Nicholas Stern, a senior British civil servant and former vice-president of the World Bank, was conceived as the definitive statement on the cost of climate change. So far, though, its legacy has been a debate among economists that has as much to do with ethics as money.

"The review and the critiques that have followed it highlight that many of the important questions boil down to what we choose to value," says Roger Pielke Jr, an expert in climate-change policy at the University of Colorado, Boulder. Perhaps the most vexing question is how current generations should value their successors' welfare.

The Stern report, published in October 2006, concluded that doing nothing about climate change would mean a long-term loss in average world consumption of 5–20% per year, whereas stabilizing greenhouse-gas concentrations at roughly double pre-industrial levels would cost 1% of global gross domestic product (GDP) by the middle of this century, with a margin of error of plus or minus 3%.

Nicholas Stern has been criticized for overestimating the current value of future generations' welfare.

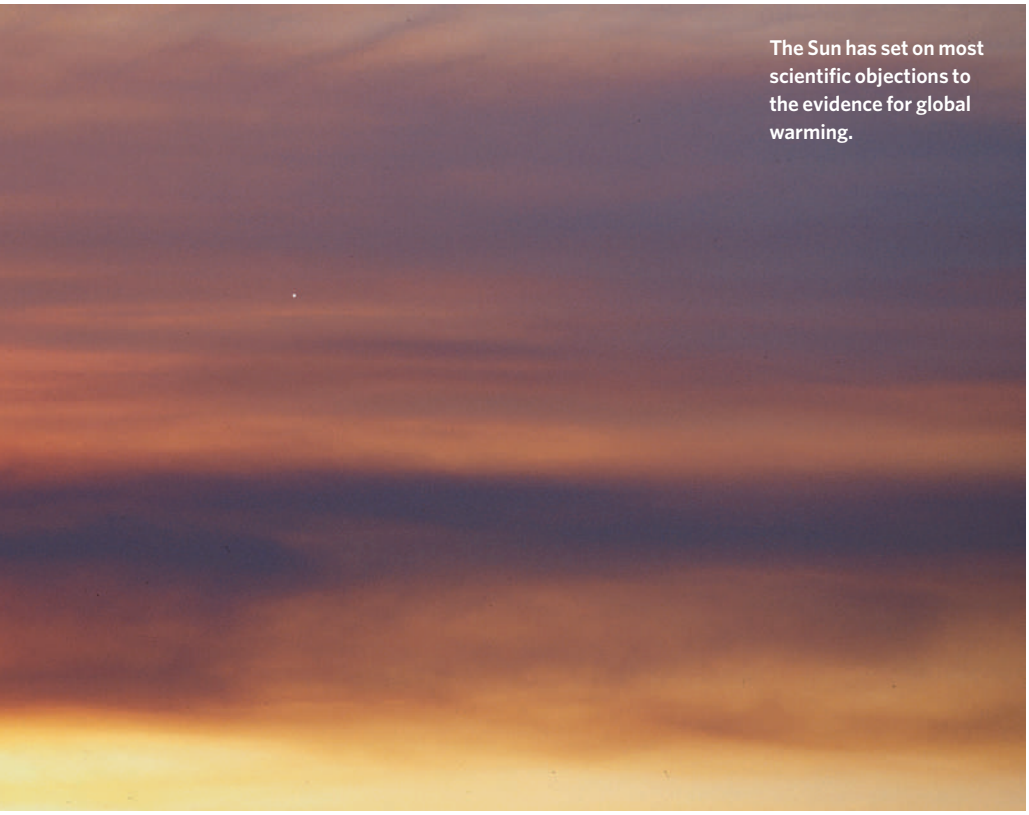
A. GRANT/AP

These calculations depend partly on the value we in the present assign to money received or spent in the future — determined by discounting. To stop climate change, we are being asked to pay now and live later, but most people value future benefits less than current costs, and will only invest if the projected pay-off is large enough.

The contentious issue is the size of the 'pure-time' discount rate, which determines how much the the welfare of current generations is valued compared with future, as yet non-existent generations.

NASA

JOIN THE DISCUSSION
Read and post comments about the IPCC report.
http://blogs.nature.com/news/blog/2007/02/climate_report.html



S. T. SMITH/CORBIS

The Sun has set on most scientific objections to the evidence for global warming.

analysed incorrectly, and when corrected do indeed show a slight warming.

Other favourite arguments of climate sceptics are also dismantled in the latest IPCC report. Urban heat islands — the fact that cities tend to heat the air above them — do exist, the report says, but have a negligible effect on global temperatures. And solar variability — natural fluctuations in the amount of the Sun's radiation reaching Earth — does affect climate, but to a far smaller extent than the burning of fossil fuels.

With less to argue about on the scientific front, climate sceptics have been turning their attention to the economics of adapting to a changing climate.

Christy believes that fostering innovation is the way to decrease reliance on fossil fuels. "We're going to look back in a century and say 'wasn't it quaint, we burned carbon'," he says. "I'm very optimistic; I see the wealth of the Earth continuing to rise. But suppressing energy is not the way. Keep energy inexpensive and affordable and allow people to do research." As for the Kyoto Protocol on climate change, the international agreement to reduce carbon emissions, Christy calls it "sinister".

Economic arguments also play a strong role in the views of sceptic Patrick Michaels, an environmental scientist at the University of Virginia who argues that taking action on climate change can have dire economic consequences. He sees the current US move towards embracing biofuel as causing corn prices to surge, triggering inflation and leaving many poor people, particularly in Mexico, struggling to buy food. "Small changes in policy can lead to a recession," he says.

It remains to be seen whether these arguments will gain much traction. Alan Thorpe of Britain's Natural Environment Research Council, which hosted an online debate to canvass climate sceptics on their views, says that such views range from lazy to devious.

"I think there is a degree to which there is mischievous use of scepticism," he says. "Sceptics want to accuse scientific society of wanting a particular policy outcome, but actual policy is up to governments." ■

Michael Hopkin

The Stern review used a very low pure-time discount rate, treating current and future generations equally. But there is little consensus on this among economists. Some, such as William Nordhaus of Yale University in New Haven, Connecticut, find low rates hard to accept, adding that the way markets work supports this position.

Stern's conclusion "depends decisively on the assumption of a near-zero discount rate", Nordhaus wrote in a critique posted on his website in November. Nordhaus recalculated Stern's assessment of the optimal rate at which emissions should be reduced using a pure-time discount rate of 3% that declines to 1% in 300 years, rather than the 0.1% rate used in the Stern review. The change cut the optimal rate of emissions reduction by 2050 from 25% to 14%. Stern retorts that a pure-time discount rate of 1% is "outrageously high", as it values things happening in 100 years as only one-third as important as things happening now.

Economist Partha Dasgupta of the University of Cambridge, UK, although largely supportive of Stern's conclusions, takes issue with the way that the report's egalitarian approach to different generations — valuing their well-being equally — is not carried through to its discussion of inequalities in today's world. Stern's calculations, Dasgupta argues, understate the degree to which a given dollar benefits a poor person more than a rich one.

Such arguments may seem like a good way to stall progress, but many, including Stern, value the attention to ethics. "How important a problem climate change is and what society should do about it is all about subjective judgements," says Mike Hulme, director of the Tyndall Centre for Climate Change Research in Norwich, UK.

"How important a problem climate change is and what society should do about it is all about subjective judgements."

However, the focus on the discount rate's effects on the estimated cost of damage has partly overshadowed the Stern report's second striking conclusion — that the worst effects of climate change can be avoided for at most 4% of GDP.

A few weeks after the main report's publication, the Stern review team released a sensitivity analysis that looked at how the conclusions might differ with more conservative estimates of damage costs derived from different value systems. Further analyses will be published in the next couple of weeks. These have shown that the conclusion that prevention will be much cheaper than dealing with the damage is fairly robust, given moderate changes in the parameters.

The debate about discount rates only looks at half the picture, says Ottmar Edenhofer, an economist at the Potsdam Institute for Climate Impact Research in Germany. Just as crucial is the debate about how to reduce emissions. ■

Lucy Odling-Smee

BUSINESS

Carbon copies

The corridors of power in the United States are ringing with the phrase 'cap and trade'. But when will carbon markets arrive in America — and what will they look like? **Emma Marris** investigates.

The trading scheme for carbon emissions started in 2005 by the European Union (EU) is far from perfect. The price of credits in its first phase, for example, bombed last spring, to the alarm of some investors — although trade in the second phase remains robust (see Market Watch, right).

But the 25-nation scheme is the best model there is — or at least that's the consensus of the people looking to build a corresponding market in the United States. Many of them say the overall approach will be most effective if the two markets can work closely together.

So pleas for such linkage were high on the agenda at a conference on North American carbon markets last month in Washington DC. "Maybe carbon markets are returning to where they were born — in the United States," Per-Otto Wold, chief executive of Oslo consultancy Point Carbon, told the meeting. "I hope that we will see compatibility between the markets."

The approach of limiting pollution by setting caps on emissions and then trading those emissions back and forth was indeed devised largely by US economists in the 1980s. Back then, they were reacting to concerns about power stations emitting sulphur and nitrogen oxides that were causing acid rain. The plan is widely seen as a success: emissions of sulphur dioxide from facilities covered by the US scheme, for example, have fallen from 16 million tonnes in 1980 to about 9 million tonnes in 2005. And US negotiators energetically advocated carbon trading in the run up to the 1997 signing of the Kyoto protocol on climate change — from which the US government later withdrew.

Now there is growing acceptance that some arrangement to cap and trade greenhouse-gas emissions in the United States is on the way. More than 80% of North American power-company managers think that 'binding carbon mandates' will come into effect sometime between 2008 and 2017, according to a survey by Cambridge Energy Research Associates of Massachusetts.

Eileen Claussen, head of the Pew Center on Global Climate Change in Washington DC, estimates that mandatory greenhouse-gas caps are "plausible by 2008 and likely by 2010". The sense that this is so has hardened considerably since the Democrats took control of Congress last November.

Industry in the United States is expecting mandatory carbon caps within the next decade.

N. UT/AP

"There are no good reasons for the US and European emission-trading schemes to differ."

Several states, including California, are also moving to implement schemes of their own. The Regional Greenhouse Gas Initiative (RGGI), for example, involves nine northeastern states and will cap carbon emissions from power plants from 2009. According to James Werner, the initiative's coordinator in Delaware, the project is based on the supposition that US-wide caps will follow. "We expect that a federal programme is coming," he says.

Industrial resistance to the idea is eroding. On 22 January, ten blue-chip companies, including Caterpillar, the construction-equipment maker, and Pacific Gas and Electric, joined forces with moderate environmental groups to call on Congress to enact legislation that would install a cap-and-trade scheme.

But major questions remain about the structure of such a scheme. One of the most basic is whether carbon credits should be auctioned off by the government — to raise revenue for its own emission-reduction efforts — or be given away free. Environmental groups tend to favour the former approach, businesses the latter. In the EU, only 5% of credits were auctioned off in phase one, but in the RGGI, all states must auction off at least 25% of their credits, and some are planning to auction off all of them. Massachusetts, for example, expects to collect between $25 million and $125 million a year by auctioning off all of its credits.

If all or part of the allocation is given away, the credit begins its life without a price attached. The price is then determined by traders' best guess of what the credits will be worth in the future. "A free allocation system has made the market less efficient," says Louis Redshaw, environmental markets chief at Barclays Capital in London. "It doesn't give a price signal."

Trading places

Another hot topic is the extent to which emissions above a cap can be dealt with by buying offsets, rather than by buying emission allocations from another company. Offsets are emission-reduction measures elsewhere in the world, many of them supported by an element of the Kyoto protocol called the Clean Development Mechanism (CDM). Last year alone, CDM contracts worth $5.6 billion changed hands. A Senate climate bill backed by John McCain (Republican, Arizona) and Joseph

Lieberman (Independent, Connecticut) would allow companies to use the CDM to keep their emissions under the cap. But critics see this as a 'dodge' that won't reduce global emissions.

Then there's the issue of which sources and gases to include. The EU scheme applies only to large industrial emitters — and covers only carbon dioxide emissions. Including the vast emissions generated by transportation is problematic. "You can't actually measure all those sources very accurately," says Damien Meadows, deputy head of the European Commission's trading-scheme unit. "Maritime and transport may come in later — but you need absolute accuracy."

Unlike the EU scheme, the McCain–Lieberman bill — and a rival proposal backed by Senator Jeff Bingaman (Democrat, New Mexico) — would include not just carbon dioxide but all six major greenhouse gases. In the sole carbon-emission market already functioning in the United States, the Chicago Climate Exchange (CCX), participants volunteer to sign legally binding contracts committing them to cut emissions of the six major gases by 6% from a pre-agreed baseline by 2010.

But because it is voluntary, the CCX is tiny, trading just 1.23 million tonnes of carbon in 2005, according to Point Carbon, compared with the 362 million tonnes traded in the EU scheme that year. The Chicago market's supporters nonetheless advocate it as the most sensible forum for emerging US trading based on mandatory caps. They say that an experimental trade conducted by Baxter Healthcare of Deerfield, Illinois, last year showed that it would make for good interoperability between US and European markets.

But Henrik Hasselknippe of Point Carbon says that carbon trading could end up happening in any forum that proves to be convenient. "The market tends to make its own marketplace," he says.

Overall, the highest concern of those at the Washington meeting was that the US approach should be tied as closely as possible to the EU scheme. That will require it to contain certain features, such as synchronized phases, and to exclude others that have been mooted in the United States, such as a cap on how high the price of emission credits can go. "A safety valve or a price cap in the United States would make linking incompatible," says Meadows. "I would emphasize the need for simplicity."

There are no good reasons for the two schemes to differ, says James Cameron, vice-chairman of European investment banking group Climate Change Capital. He calls for the US system to mirror the European one. "Don't reinvent, connect," he says. ■

See Commentary, page 595.

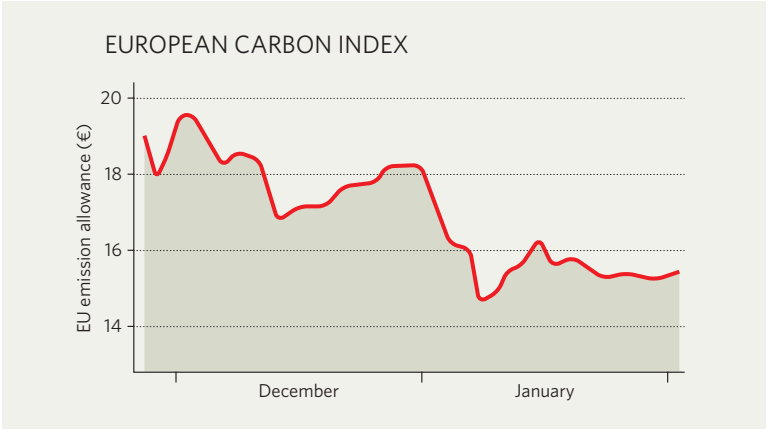
IN BRIEF

SAFETY STEPS **The US Food and Drug Administration (FDA) has announced a set of measures to improve the science it does, the safety of drugs it has approved, and the way it communicates safety concerns to the public (see *Nature* 443, 372; 2006). Among these is a pilot programme to assess the safety of drugs with novel active ingredients 18 months after they have gone on market. But Senator Christopher Dodd (Democrat, Connecticut) says the measures don't go far enough and, with Senator Chuck Grassley (Republican, Iowa), he has introduced legislation that would give the drug agency additional powers to enforce the safety of existing drugs.**

CELL BANK **Richard Branson, flamboyant chairman of Virgin Group, has set up a company that will charge parents £1,500 (US$2,900) to store blood stem cells from umbilical cord tissue for possible future use in the treatment of blood diseases. Virgin Health Bank, which will run the store in collaboration with London-based investment company Merlin Biosciences, will also put half of each sample in a bank for public health use. Branson says that profits will be donated to fund stem-cell research.**

BIOFUELS WINNER **BP says it will spend $500 million over ten years on biofuels research at an Energy Biosciences Institute to be established at the University of California, Berkeley. The university won the race to host the centre against rivals reportedly including the Massachusetts Institute of Technology and Imperial College London. Its bid was backed by Arnold Schwarzenegger, the governor of the state, which will also contribute $40 million to the centre. It will operate as a consortium that will include the nearby Lawrence Berkeley National Laboratory and the University of Illinois at Urbana-Champaign.**

MARKET WATCH



SOURCE: EEX

Allowances to emit one extra tonne of carbon dioxide in Europe during 2007 can now be bought for a song — €2.30 (US$2.99) last week — and all eyes are turning towards the second phase of the European Union (EU) emissions-trading system, which runs from 2008 to 2012.

Carbon futures for 2008 are being traded at €15 — down from almost €20 two months ago and sharply below last April's peak of €32.

Market-watchers say that the decline reflects steadily falling prices for natural gas. The resultant switch from coal to cheaper and cleaner gas by European electrical utilities has dampened demand for emissions credits, as companies grow confident that they will stay below their emissions caps. And on 30 November, the European Commission loosened restrictions on the number of credits that companies can obtain on the basis of emission reduction projects outside the EU.

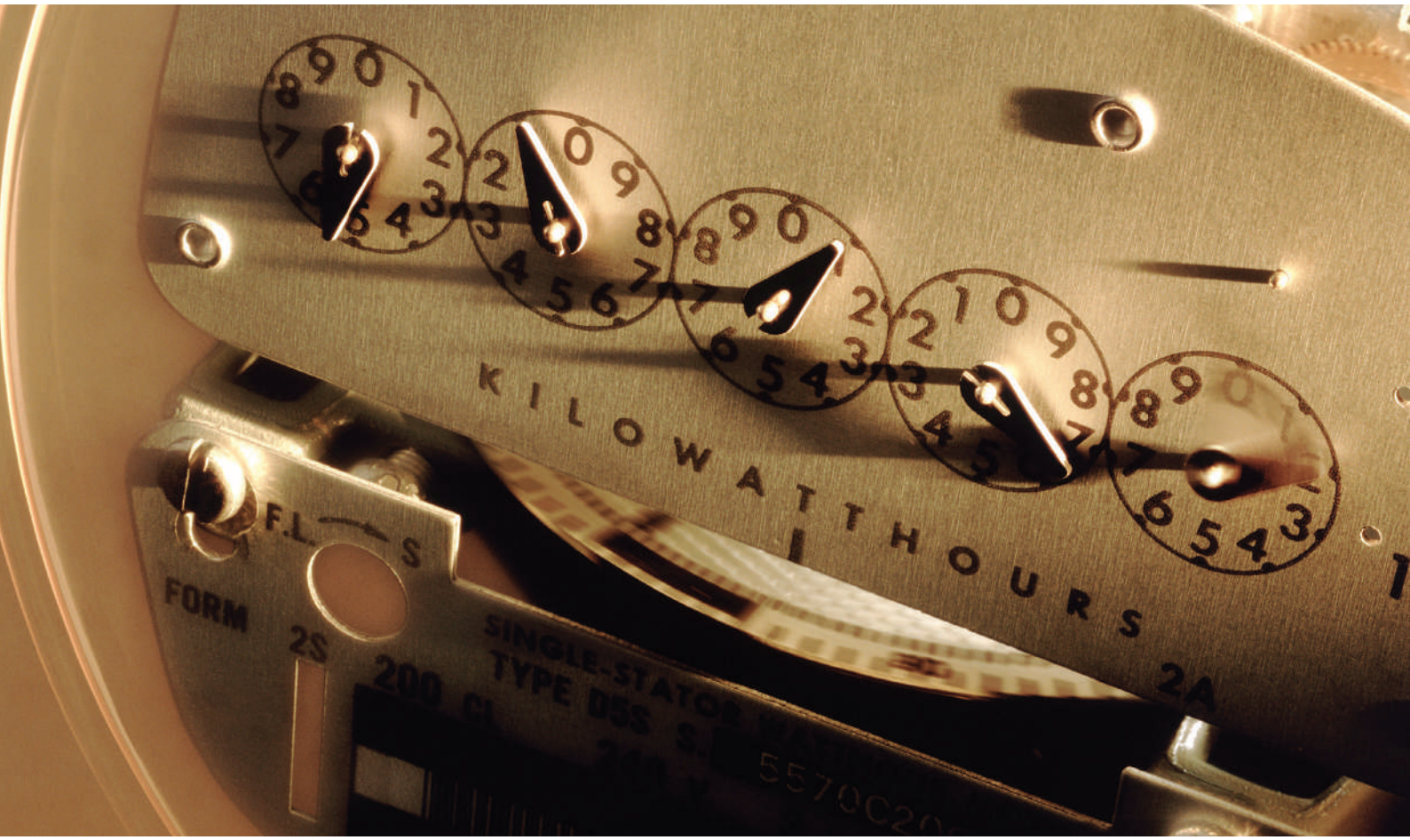
As a result, phase-two allowances are selling for a lot less than the €20–30 predicted by London-based consultancy, New Carbon Finance, back in November.

But Guy Turner, an analyst at the consultancy, says that right now it is more important for the market to prove that it can operate successfully, than for it to attain a high price that will force emissions reductions in Europe. "Political integrity is more important than environmental integrity at the moment," he says. "The EU needs to show the United States and China that it's possible to manage the process — that emission trading can be done and is a club worth joining." ■

Quirin Schiermeier

SUPER SAVERS

Energy efficiency is one of the least flashy but most promising ways to cut carbon dioxide emissions. In the first of two features, **Declan Butler** explores the energy-saving possibilities of an intelligent electrical grid. In the second, **Zoë Corbyn** looks at how labs can cut their energy use.



PHIL BANKO/CORBIS

Meters to manage the future

"Switch off the washing machine or dryer for the next 3 minutes, and let me buy, at 10 cents per kilowatt hour, 20% of your solar energy output." Such instant electronic transactions between electricity distributors and smart electric meters in millions of homes and businesses are set to add some badly needed intelligence to the electricity grid, bringing greater efficiency and reliability.

Ways that humanity can adapt their use of energy to the realities of global warming tend to focus either on supply — renewable technologies such as solar power — or demand, personal abstinence of various sorts. But there is untapped potential in the space between supply and demand — the mundane world of energy distribution. Making the electrical grid more efficient would offer benefits on the demand side, by helping users to consume less energy, and in terms of supply, by providing better ways to handle the intermittent and distributed nature of alternative energy sources.

The humble electric meter might seem an unlikely place to start a revolution. But today's centralized grid, which has electricity fanning out from a few large generators across transmission lines to end users, is mostly dumb. Utility companies have detailed data on events at the power plant or on their transmission line, but once the electricity radiates out to the community they have no idea where it goes, let alone how to manage it.

Smart meters change that by providing data on what is happening in every corner of the grid at any instant. Utilities in Europe and the United States are rolling out millions of meters that send real-time data on the electricity use of individual homes and businesses via the Internet, or along the electricity supply. In turn, these meters can receive real-time data on grid conditions, load and pricing.

The meters are a stepping stone to smart grids, which would provide a modern, distributed network of computing and telecommunications — a dynamic 'energy Internet'. And just as the Internet triggered an explosion of innovative technologies and services, not least the Web, new and more efficient electricity services will flow from smart grids.

"Smart meters are the harbinger of the future," says Steven Hauser, vice-president of strategy at GridPoint, the first company to sell such energy management systems direct to

M. KARRASS/CORBIS; P. DA SILVA/CORBIS

consumers. "They are the breakthrough that has helped utilities understand that you can see what's happening on the network right down to the customer level," he says.

Enel, an Italian electricity company based in Rome, is the world's biggest user of smart meters. It has installed more than 30 million meters since 2001, mostly to improve billing. In the United States, the world's largest consumer of electricity, California is one of many states following suit. By 2011, Pacific Gas & Electric, which serves northern and central California, will have supplied its 9.3 million customers with smart meters, and Southern California Edison is rolling out 5 million.

The latest meters are built to meet smart grid standards designed to ensure that all networks and devices speak the same language — the grid equivalents of the data-transfer protocols that made the Internet possible. The standards are being driven by a swarm of research efforts, including the US Department of Energy's (DoE) GridWise programme, and the European Union's SmartGrid initiative.

On the supply side of the efficiency ledger, smarter devices and richer data will facilitate the development of sophisticated energy-management software. A smart grid can accommodate a greater diversity of fuel supplies and, in particular, intermittent energy sources, such as wind and the Sun, says Kevin Kolevar, director of the DoE's Office of Electricity and Energy Assurance.

"Smart meters are the harbinger of the future."
— Steven Hauser

A centralized grid is inefficient and costly. Only a third of the fuel energy burnt in power plants ends up as electricity, with half lost as waste heat, and a further 8% lost along long-distance transmission lines. Moreover, 20% of generating capacity exists only to meet peak demand, so it runs just 5% of the time and provides just 1% of supply. The grid is often congested because it relies on a few high-traffic arteries. The congestion amplifies the inefficiency because if the utility cannot redirect power from efficient sources, they have to turn to costlier, dirtier and more inefficient sources to meet peak demand.

A more distributed grid, by its very architecture, can improve efficiency by matching local supply with demand. With multiple decentralized energy sources, electricity can be generated close to the point of use, avoiding the losses and congestion that result from long-distance transmission. Some of the most efficient energy sources are small turbines powered by natural gas, or biogas, which use waste heat to provide heat and hot water to the local area, and convert energy with 70–85% efficiency.

The world leader in decentralized power, Denmark, now generates half its electricity through decentralized grids, with combined heat and power accounting for 80% of local-area heating, and wind power about 20% of all electricity. As a result, its carbon dioxide emissions have tumbled from 937 grams per kilowatt hour in 1990, to 517 grams per kilowatt hour in 2005. Denmark began its push towards decentralization two decades ago, before smart-grid technology was available, relying mostly on tough regulations to force the change.

The key to Denmark's success was standing up to the utility companies, says Henrik Lund, an expert in energy systems at Aalborg University. For instance, the government required that energy companies buy back electricity from consumers at 85% of the price. Lund recalls being told by the utilities in the 1980s that it was not technically possible for more than 10% of electricity to be provided by wind power. The country wouldn't be where it is now, he says, if they had taken statements from the utility companies as fact.

This is an important lesson for the smart-grid movement. Although the infrastructure for intelligent grids is falling into place, overcoming institutional and market barriers remains a major issue, says Kolevar. "Many utilities have a disincentive to push distributed generation, as generally it is the customers who are the biggest beneficiaries, with the utilities seeing lower sales."

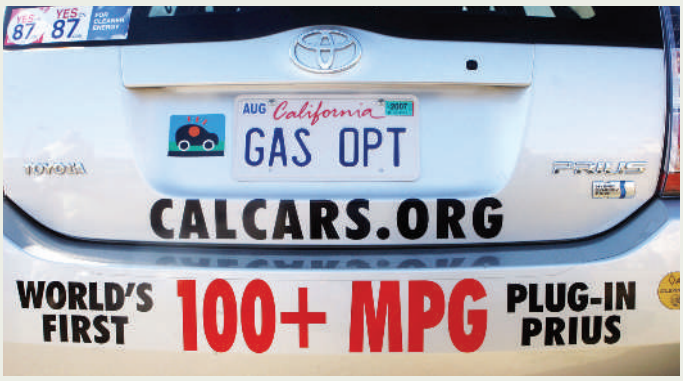
Resistance to distributed generation can be seen in the development of 'net-metering' laws, which oblige US utilities to allow consumers to sell energy back to the grid. Forty states have implemented such laws, but by 2004, just 15,200 US customers were taking advantage of the rules, 13,000 of whom were in California, according to the Network for New Energy Choices, a New York-based advocacy group.

In a survey the group published in November, the fingerprints of utility companies were found in nearly all the state laws. Indiana forbids commercial and industrial companies from taking part, for example, whereas Arkansas pays such poor rates for consumer-generated electricity that only three customers have participated since its launch in 2001. California emerged among the best, but even here the state has capped solar-energy generation at 2.5% of utilities' peak demand, arguing that it needs to assess its impact on the grid.

But the regulatory environment continues to change. The 2005 Energy Act requires that all federal buildings be equipped with two-way metering and energy-management systems by 2012, creating a huge potential market. Eventually, all states will be expected to offer net-metering and time-based pricing. And with the traditional grid straining under peak load, utility companies are waking up to distributed generation and storage as being vital to meet future demand, says Kolevar.

Companies are also emerging to exploit the new regulations. GridPoint's device, for example, can manage a residential solar panel or windmill, store generated energy in a battery that holds 12 kilowatt hours of electricity,

Plug in, turn on...sell out



K. KENNELL/ZUMA PRESS/NEWSCOM

"Forget hydrogen, forget hydrogen, forget hydrogen." James Woolsey, a former CIA director, reckons that there's a faster and cheaper way to solve America's energy problem — plug-in electric cars, acting as a massive distributed battery for the electricity grid.

Woolsey, founder of Set America Free, an organization based in Washington DC that advocates the reduction of US dependence on oil, was speaking in January 2006 at the launch of Plug-In Partners, an advocacy group led by the city of Austin, Texas. Its members have already pledged to buy almost 10,000 cars for municipal fleets.

The idea is to replace millions of gas guzzlers with electric vehicles that are powered by off-peak mains electricity. Millions of vehicles could be charged without building any new power stations, as existing plants rarely run at full capacity. Part of the stored energy could then be sold back to the grid at periods of peak demand.

Integrating vehicles and the electricity grid, the two largest consumers of energy, makes a lot of sense, according to a study published last month by the Pacific Northwest National Laboratory. It concluded that the existing grid capacity could power 217 million light-duty vehicles (three-quarters of the light-duty fleet).

Moreover, because the light-duty vehicles consume 97% of the petrol supply, switching to plug-in hybrid electric vehicles (PHEVs) could save 6.5 million barrels of oil a day — or half of US imports. The scheme would require running power plants at higher constant levels, but because they are more efficient than car engines, the net balance would be a 27% reduction in greenhouse-gas emissions.

PHEVs are similar to today's electric car, but they have bigger and more expensive batteries to cover average daily driving needs (53 kilometres), with a fuel engine kicking in for longer trips. Assuming that PHEVs would cost US$6,000 to $10,000 more than a regular car, the study estimated a user would recover the investment in 5–8 years. Those economics would improve further once batteries could be mass produced, says Steven Letendre, an expert in vehicle-to-grid economics at Green Mountain College in Vermont.

The snag is that for now, very few PHEVs are available. Only one small company, AC Propulsion in California, currently makes them, although General Motors, Toyota and DaimlerChrysler, are working on prototypes. "The real key to making PHEVs work is the vehicle — they must be attractive in terms of performance and cost to the US consumer — or the high penetrations we studied will not come to pass," says Robert Pratt, one of the authors of the study. A 24 January executive order by US President George W. Bush may give PHEVs a federal boost, by requiring US agencies to adopt them as soon as their life-cycle costs become comparable to those of gas-fuelled vehicles.

Without such incentives, or with continued hikes in oil prices, Pratt reckons that it could take 20 years before PHEVs make a serious dent in the US vehicle fleet. Woolsey thinks that they still make much more sense than the hydrogen economy, which requires changing the entire energy and transport structure. By comparison, he said, PHEVs demand "a bigger battery, and yes, an infrastructure investment: an extension cord. Every family would need an extension cord." **D.B.**

and then buy and sell to the grid depending on price. The battery also provides 8–12 hours of backup supply for a typical home.

The units contain no one technological breakthrough, admits Hauser, but instead pull existing devices together into a single box and add clever software. The climate is now ripe for such units, he says, owing to increasingly friendly legislation. Gridpoint units are not cheap, ranging from $6,000 to $16,000, but buyers will benefit from a 30% income tax credit on the purchase price through the 2005 Energy Act (with a $2,000 cap for residential customers), and could shave 10–15% off their electricity bills.

"Many utilities have a disincentive to push distributed generation."
— Kevin Kolevar

With appropriate regulation, real-time pricing should increase the competitiveness of renewable energies and of storage because local producers could sell power when prices are highest during peak demand and purchase it at the lower rates during off-peak periods. The overall effect would be to flatten out the load and price curve, and shave demand at peak hours, says Kolevar. The extra capacity offered by a distributed market should also mean that fewer of the inefficient central generators would need to be built.

Capacity demands can also be buffered by energy storage systems. A suite of distributed storage technologies are available, including fuel cells, flywheels, superconducting magnetic energy storage and advanced batteries, but they have never been used or tested on a large scale. Smart grids should change the economics in their favour.

Electricity storage is another area in which GridPoint is trying to turn theory into practice. It has signed a deal with the power company Cogentrix, headquartered in North Carolina, and is in discussion with others to test thousands of its units as a backup supply to ease demand on the grid at peak periods.

Even the idea of using plug-in hybrid vehicles as distributed energy storage is now being taken seriously, says Kolevar (see 'Plug in, turn on...sell out'). It was previously considered a research area, but Kolevar sees real interest in the idea, adding that what is now needed is a "large and definitive" demonstration project .

Another hoped-for result of real-time pricing is that homeowners and businesses will reduce their energy consumption. Typical wholesale electricity prices vary wildly, even over the course of a day, but consumers are usually charged a flat rate calculated as an average estimate over months. Smart meters can, however, make decisions that are based on real-time pricing information and the user's preset choices, such as to turn down the air conditioning if the price goes above a certain level.

Market research says that introducing such price information will encourage consumers to use less electricity overall, not just at peak periods, although these ideas still need to be tested in the marketplace. The Pacific Northwest National Laboratory in Richland, Washington, is testing a GridWise system to send real-time pricing via the Internet every 5 minutes to 200 homes on Washington's Olympic Peninsula. Heating and use of other appliances is adjusted automatically according to the pricing and usage preferences set by the users. Although the results are not yet in, preliminary data show that customers have responded well to the system and are highly responsive to the price, with consumption falling by more than 10% at peak times, says Donald Hammerstrom, a project manager for GridWise at the Pacific Northwest National Laboratory.

Experiments like this would have been unthinkable a few years ago, when an infrastructure for grid data and communications did not exist, says Hauser. With intelligent grids now emerging, the stage is set for a wave of innovation that will generate, store and manage electricity more efficiently, resulting in energy production and consumption that are less polluting and more cost-effective. ■

Declan Butler is a senior reporter for *Nature* based in Paris.

D. SIMONDS

Experimenting with efficiency

In 1971, no one really worried about energy efficiency — certainly not at Fermilab, the US Department of Energy's (DoE's) particle-physics laboratory in Batavia, Illinois. A new superconducting ring for the lab's accelerator, designed to push particles closer than ever to the speed of light, was enthusiastically talked up as the 'energy doubler'. A few years on, though, as the lab prepared its funding bid against the backdrop of the oil crisis, the system started being referred to as the 'energy saver', shifting the emphasis from increased performance to reduced power requirements.

When you spend US$1 million a month on electricity, as Fermilab does, such care in presentation is important. Under most circumstances, though, scientists give scant thought to totalling up the wasted power and unnecessary carbon emissions that their work generates. Geoffrey Bell, who works on reducing the energy consumption of Lawrence Berkeley National Laboratory in California, is one of the exceptions. He's eager to point out that a traditional fume cupboard, for example, uses as much energy in a year as three US households. "If you have a laboratory with 50 of those, you've made a town in one building!"

Laboratories consume between five and ten times more energy than office buildings — but they are also rarer and more diverse in design, making neat, generalized solutions to profligacy hard to find. Add that to concerns about safety and a lack of transparency in costs (few scientists know or care what their lab's electricity bill is), and you get a 'that's just the way they are' mentality. That's the mindset that the Labs21 programme, an initiative started by the DoE and the US Environmental Protection Agency (EPA), exists to challenge.

According to the EPA's Dan Amon, who oversees Labs21, the big beast for energy experts to tame is ventilation: some 60–70% of the energy a lab uses goes on moving, heating and cooling the air, with the rest split about two to one between appliances and lighting. Labs21 concentrates its efforts on 'wet labs' — chemical or biological labs fitted with fume cupboards, which need to exchange the lab's air with outside air at a high rate. Bell, a member of Labs21, says that most of these facilities have been designed according to "what is in the drawer", with fume cupboards making many more air changes, and at higher velocities, than is necessary. In the United States and Britain, safety standards require that air is exchanged 6 to 12 times an hour. Yet rates as high as 15 to 25 per hour are not unusual. Even a small reduction from 12 to 10 air changes per hour can reduce the amount of fan power by more than 40%, says Bell.

> **"The more modern the laboratory, the worse its energy consumption."**
> — Peter James

A key to progress here is replacing fume cupboards that replace air constantly with technology that has a variable air volume. A variable-air-volume fume cupboard adjusts the speed of the fan, and therefore the energy use, to the position of the cupboard's sash opening. As the requirements for ventilation change — when researchers open and close the sashes — the building's exhaust and supply fans adjust accordingly. Bell says that some US labs have used a carrot-and-stick approach to encourage careless researchers to make the most of this technology by using the sashes on their cupboards sensibly. "Beware the sash police," he jokes, "they might make you walk around with your apron on backwards as a punishment."

The new Science and Technology facility at the DoE's National Renewable Energy Laboratory (NREL), completed in July last year, is designed on Labs21 principles, with an airflow rate of at least 0.3 cubic metres per minute per square metre (see 'Look, no carbon!'). With

ceiling heights of three to four metres, this means about six air changes an hour. In addition to variable-air-volume technology, the lab has systems to reuse the heat from exhaust air and from the water used to cool the equipment. It also makes extensive use of natural light and groups activities that require high ventilation rates together so they can be dealt with differently from the rest of the space. "Overall, if you were to compare this facility with a more conventional lab, the savings are about 41% of the overall energy cost," estimates NREL's Pete Sheldon, who coordinated the lab's design. "This is about $96,000 per year." He adds that it's the quality of the light that the researchers comment on most frequently.

Sustainable inklings

According to Phil Wirdzek, who had Amon's job at the EPA in the early 1990s, the Labs21 programme started more or less by accident. When reporting their annual energy consumptions to the US Congress, the EPA and the DoE went beyond what was required and included the energy bills of their various facilities in calculations. Wirdzek was responsible for his agency's report. "It was like, 'holy mackerel did I get us into trouble!' But it started Labs21 because we began to say: well, how do you fix these things?"

Now administered in part by Wirdzek's not-for-profit organization the International Institute for Sustainable Laboratories, Labs21 runs an annual conference and trade show, offers design assistance and training to labs, which become 'partners', and has recruited 4,000 engineers, architects and manufacturers to its green agenda. Designing an energy-efficient lab requires a different way of thinking, according to Amon. "If you don't tie into the community [of people thinking differently], then you are going to do what you have always done," he says. Bell emphasizes that when designing or refitting a lab, the key is to get all concerned — lab managers, contractors, scientists, union representatives, cleaning staff — to buy into the process from the beginning, and to ask why things are the way they are. Why are there so many fume cupboards and refrigerators? Why is the temperature tolerance so tight? Why are air-flow rates not reduced at night?

Although he doesn't put it all down to Labs21, Amon guesses that about a quarter of the labs in the United States now use energy-efficient design principles. Thanks to interest from the US Green Building Council in Washington DC, labs will soon be eligible for the prestigious Leadership in Energy and Environmental Design (LEED) awards that are already in place for office buildings, schools and shops.

Peter James, a professor in environmental management at Bradford University in the United Kingdom and coordinator of a public-sector initiative to raise universities' environmental performance, is importing the Labs21 approach to Britain. According to James, who took part in an analysis of data from around 50 campuses in Britain a few years ago, "a really horrifying thing we found was that, by and large, the more modern the laboratory, the worse its energy consumption". He points to more stringent health and safety requirements in the past ten years and to the wasteful margins being built into ventilation rates; like Bell and Amon, he is eager to assure people that the waste can be cut without compromising safety.

Two-way systems

Mike Dockery, a UK-based consultant who designs laboratory systems, argues that Britain can teach the United States a thing or two — specifically on 'FlexiLab' methodology. First developed by the drug giant GlaxoSmithKline for its UK facilities, the FlexiLab system is now being rolled out at the company's sites in the United States. Portable variable-air-volume fume cupboards connect into prefabricated, standardized ducting and service 'spines'. It makes it possible to change the type of science a lab is doing over the weekend, says Dockery, adding that reusability also means less wasted equipment when the lab's function changes. James says that in Britain, public labs are playing a game of catch-up with private companies.

But are scientists really ready to embrace sustainability in their own backyards? A recent online survey by the UK Department for Environment, Food and Rural Affairs had telling results. Of 400 scientists across a wide range of disciplines questioned in August 2006, 95% agreed that science and technology were important if sustainable solutions were to be developed for the future, but only 40% said that they always or often considered the effect their work would have on the environment when planning their research. Of those who didn't think about the environmental impact of their research, 53% said it was because they felt it wasn't relevant to their area of science.

The results don't surprise Bell. "Scientists tend to be in their own little world most of the time. They are not necessarily belligerent or disrespectful, they are just very focused people…and they really don't understand how they might be knocking things over in the process." "But," he adds, "at some point in time you have to pull them out of that blinders-on attitude to say, 'Look — you did great science over here and you made this new wizard medicine — but look what you did over here.'" ■

Zoë Corbyn is a freelance science writer based in London.

This week, Geoffrey Bell will be answering questions about this subject on the *Nature* newsblog (http://tinyurl.com/3a3d4f), where you can tell us what your lab is — or isn't — doing about energy efficiency.

Look, no carbon!

Two US labs have gone further than just using efficient energy, and have become 'carbon neutral'. The National Renewable Energy Laboratory (NREL) in Golden, Colorado (see right), and the Robert Kerr Environmental Research Centre in Ada, Oklahoma, think that they are the only labs in the world to have reduced and offset their carbon emissions to zero — blazing a trail for others to follow.

NREL, which researches everything from photovoltaics to biomass energy, has just completed its first carbon-neutral year by balancing its power use and emissions generated from staff flying to conferences and commuting to work with various energy-saving initiatives and offsets.

The lab uses ethanol-fuelled vehicles and has designed or refitted its buildings to Labs21 principles (see main story). It also uses energy generated from its experiments to power the facility, although this accounts for only about 5% of the lab's needs. "If we can capture power — like on our wind experimental site — then we will use it," says Robert Westby, who oversees sustainability at the facility. But the experiments are experiments first, practical generators second.

P. CORKERY/NREL/DOE

Burning tree thinnings from the nearby forest, NREL should be able to generate as much as 20% of its own energy this year, Westby estimates, thanks to a plant being paid for, designed, installed and operated by a private company. "The private sector sees that it can make money doing this," explains Westby, "so the lab doesn't have to make the investment."

The Robert Kerr lab, which does mostly groundwater research, decided to go 'zero emission' in April 2005. Like NREL, the lab supplements energy-efficiency measures and renewable-energy sources with the purchase of green-energy certificates — though travel is not offset. The 1960s lab has replaced its former natural-gas use with a 'ground-source heat pump'. A series of wells (also run as a public–private partnership) tap into the ground's constant temperature, cooling the lab in summer and warming it in winter. A variable air-volume system also reduces the amount of air and electricity required. **Z.C.**

How important is immune memory to invertebrates?

SIR — Margaret McFall-Ngai's Essay, "Care for the community" (*Nature* **445,** 153; 2007) suggests that the unique existence of immune memory in vertebrates (the 'adaptive' immune system) could have evolved to recognize and manage beneficial microbe communities that invertebrates usually don't use.

This hypothesis is interesting and deserves consideration, although I would like to bring readers' attention to some earlier research, for example by J. Kurtz and K. Franz (*Nature* **425,** 37–38; 2003), mostly by evolutionary ecologists, reporting immune memory in invertebrates.

The vertebrate immune memory is based on immunoglobulins that invertebrates lack. To date, we know almost nothing of immune-memory mechanisms in invertebrates; hence the phenomenon has been observed before being mechanically understood. It is the opposite of the current trend in which genes are discovered before their functions are known, and is a good illustration of the importance of the complementarity of disciplines in biology.

Even so, the role and significance of immune memory in invertebrates remains unknown, and the interesting ideas discussed in this Essay could help to explain why this function might not be as central for invertebrates as it is for vertebrates.

Simon Fellous
Laboratoire de Parasitologie Evolutive, CNRS-UMR 7103, Université Pierre & Marie Curie, Paris 75005, France,
and Biology Division, Imperial College London, Silwood Park Campus, Ascot SL5 7PY, UK

Getting that first scent of life while we're in the womb

SIR — In the opening sentence of his excellent Brief Communication "Underwater 'sniffing' by semi-aquatic mammals" (*Nature* **444,** 1024; 2006), Kenneth C. Catania states that mammals cannot smell underwater because it is impossible to inspire air. It is true that there is no air underwater; however, there has been a long debate about whether air is actually necessary to smell.

Ernst Heinrich Weber, a German physician who pioneered experimental psychology in the nineteenth century, heroically filled his nostrils with eau-de-Cologne diluted in water, and reported that he could not perceive the distinct smell of the dilution. Weber concluded that odours can only be smelled in air.

This was the reigning wisdom for the next 40 years, until Eduard Aronsohn repeated Weber's experiment (*Arch. Physiol.* 321–357; 1886). Aronsohn reported a "horrible explosion of the most unpleasant and painful sensations in the nose" after filling it with diluted eau-de-Cologne. He learned from the experience, and from then on used a warm sodium chloride solution instead of cold water.

Aronsohn continued to do experiments on himself — and on colleagues and friends — with clove oil, camphor, eau-de-Cologne, coumarin and vanillin. He came to the conclusion that all odours could be smelled when he filled his nose with a dilution of each one in salt water.

Of course, this ability won't help humans to follow an earthworm scent trail in a river, as it does the ingenious star-nosed mole. But at least it allows us and other mammals to smell in the absence of air in the womb (B. Schaal, L. Marlier and R. Soussignan *Chemical Senses* **25,** 729–737; 2000).

Andreas Keller
Laboratory of Neurogenetics and Behavior, Rockefeller University, 1230 York Avenue, Box 63, New York, New York 10021, USA

Colour-blindness: how to alienate a grant reviewer

SIR — With regard to recent Correspondence (*Nature* **445,** 147 & 364; 2006) on the prevalence of scientific figures that are difficult for people with red–green colour-blindness to read, I am compelled to support Chris Miall's position.

As a red-green colour–blind (deuteranope) scientist and graphic designer, I have long campaigned for figures to be accessible to an entire audience. I do so, in part, by leading seminars training my colleagues to create accessible figures.

One of the key resources I employ in this crusade is a website by Masataka Okabe and Kei Ito: http://jfly.iam.u-tokyo.ac.jp/html/color_blind.

I strongly urge all authors to visit this site, which both describes the need for creating accessible images (including simulations of colour-blindness for those who are curious) and, more importantly, provides instructions for making figures comprehensible to everyone. This includes instructions on how to pseudo-colour images containing red and green fluorescent signals — one of the most hated types of graphic among people with colour-blindness. Authors will find it is surprisingly easy to accommodate the colour-blind when creating figures.

Anyone who needs to be convinced that making scientific images more accessible is a worthwhile task should consider that colour-blindness is common, affecting 5–10% of males. If your next grant or manuscript submission contains colour figures, what if some of your reviewers are colour-blind? Will they be able to appreciate your figures? Considering the competition for funding and for publication, can you afford the possibility of frustrating your audience? The solution is at hand.

Joseph A. Ross
Peichel Laboratory, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Mailstop D4-100, Seattle, Washington 98109, USA

Sherlock Holmes's skills as a philosopher? Elementary

SIR — Richard Gregory, in Books & Arts, is not the only one to find professional inspiration in Arthur Conan Doyle's fictional hero Sherlock Holmes ("The great detective" *Nature* **445,** 152; 2006).

See, for example, the work of Umberto Eco and Thomas A. Sebeok, comparing the reasoning methods of Holmes and of Edgar Allan Poe's detective, C. Auguste Dupin, with those of the logician Charles Peirce, in *The Sign of Three: Dupin, Holmes, Peirce (Advances in Semiotics)* (U. Eco and T. A. Sebeok, Indiana Univ. Press, 1984).

One might wish to follow Holmes's example with caution, however. As ably documented by Dr Watson in *A Study in Scarlet*, Holmes's scientific credentials are mixed. Watson's note, headed "Sherlock Holmes: his limits", includes:

> *"[knowledge of] Astronomy: nil… Botany: variable. Well up in belladonna, opium, and poisons generally. Knows nothing of practical gardening. Knowledge of geology: practical, but limited. Tells at a glance different soils from each other. After walks has shown me splashes upon his trousers, and told me by their colour and consistence in what part of London he had received them. Knowledge of chemistry: profound… Anatomy: accurate, but unsystematic."*

Famously, despite referring to his methods as "the science of deduction and analysis", Holmes was unable to distinguish between a deductive and an inductive inference. This failing might be accounted for by the fact that Watson also documented Holmes's knowledge of philosophy as "nil".

Philip Beaman
School of Psychology and Clinical Language Sciences, University of Reading, Earley Gate, Whiteknights, Reading RG6 6AL, UK

Contributions to Correspondence may be submitted to corres@nature.com. They should be no longer than 500 words, and ideally shorter. Published contributions are edited.

COMMENTARY

Is the global carbon market working?

The Clean Development Mechanism can be viewed not only as a market, but also as a subsidy and a political mechanism. **Michael Wara** argues that it has been most effective, so far, in achieving its political goals.

A perennial problem in international climate politics is how to engage developing nations in controlling greenhouse-gas emissions. These countries have more immediate priorities than climate change. Yet they must be part of any effective solution to global warming, for their emissions are high and rising (although not nearly as high, on a per capita basis, as those of the industrialized world). To encourage developing-country participation, the Kyoto Protocol established a global market for emissions reductions in 2003 called the Clean Development Mechanism (CDM). This market is now mature enough for analysis of its successes and shortcomings.

The CDM works by paying developing countries to adopt lower-polluting technologies than they otherwise would. For example, rather than building an inefficient but cheap coal-fired power plant, a Chinese utility might choose instead to build a more efficient gas-fired plant that emits less carbon dioxide. The difference in potential carbon emissions between the coal and gas plants can, after monitoring and certification, be converted into CDM credits that can be sold to an industrialized nation party to the Kyoto Protocol. The revenue from the credits enables the utility to afford the more expensive gas plant. The purchase of low-cost credits by industrialized nations to offset their own emissions reduces the cost of complying with Kyoto. The mechanism works because it is cheaper to construct low-carbon energy infrastructure from scratch in developing nations than to modify or replace existing technology in industrialized nations.

The CDM has become an important component of how European governments intend to comply with their Kyoto commitments because it reduces the cost of compliance. It is also essential to energy companies and others involved in the European cap-and-trade programme for CO_2, called the Emissions Trading Scheme (ETS). Last year, the United Kingdom proposed that ETS emitters with CO_2 caps should be allowed to use CDM credits to meet up to two-thirds of their ETS effort[1]. Together, the CDM and ETS are the keystones of an emerging global regime of linked but distinct markets for greenhouse-gas emission controls[2].

G. BAKER/AP

The carbon market has not yet convinced China to switch from coal to gas.

"Developing nations must be part of any effective solution to global warming."

But is the CDM working? The answer depends strongly on the criteria against which its success is evaluated. There is near unanimous agreement that the CDM has succeeded in engaging many buyers and sellers and substantially reducing emissions of the six Kyoto Protocol gases (CO_2, methane, nitrous oxide, hydrofluorocarbons, perfluorocarbons and sulphur hexafluoride). So far in the CDM scheme, the projected reductions for all these gases combined add up to more than 1.75 billion tonnes of CO_2-equivalent emissions[3]. This equates to annual reductions of 278 million tonnes, a very small fraction of the annual global CO_2 emissions (26 billion tonnes in 2003)[3].

Active primary and secondary (resale) markets in CDM credits have emerged, along with sophisticated systems for verification and delivery. Developing nations that were initially sceptical of the CDM — notably China and India — have entered the market with great enthusiasm and now sell the most credits. The regulatory regime administered by the United Nations has overcome both funding and logistical hurdles to emerge as a relatively successful arbiter of the global marketplace. These political accomplishments are outstanding, but they are not sufficient to judge the effort a success.

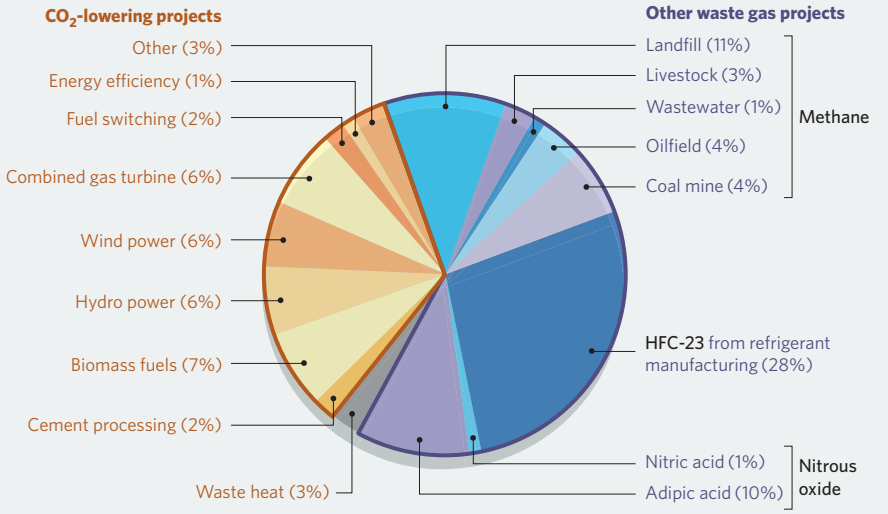
In other, and perhaps more important ways, the CDM is failing to deliver results. Initially, the market was expected to create strong incentives to invest in infrastructure for low-carbon energy in developing countries. Although many gases cause global warming, CO_2 matters most because it is emitted in prodigious quantities and has a long atmospheric lifetime. The energy sector is generally the largest emitter of CO_2 in any country. Yet a detailed look at CDM projects producing and selling credits reveals that nearly two-thirds of emissions reductions involve neither CO_2 nor energy production (see chart, overleaf)[4].

Cashing in

The largest volume of credits, almost 30% of the entire market, come from capturing and destroying trifluoromethane (HFC-23), a potent greenhouse gas that is a by-product of the manufacture of refrigerant gases. At current carbon market prices (~€10 (~US$13) per tonne of CO_2) and neglecting taxes, these HFC-23 credits amount to €4.7 billion up to 2012 (the end of the first compliance period

EXISTING PROJECTS IN THE GLOBAL CARBON MARKET

Breakdown of 1,534 projects in the CDM pipeline.



of the Kyoto Protocol). In fact, HFC-23 emitters can earn almost twice as much from CDM credits as they can from selling refrigerant gases — by any measure a major distortion of the market[4]. The distortion exists because it is extremely cheap to cut HFC-23 emissions from these facilities. Indeed, in the industrialized world similar manufacturers have chosen to reduce their emissions voluntarily. An alternative approach to cutting HFC-23 emissions from the small number of refrigerant producers in the developing world (17 at the last count) would be to pay them for the extra cost of installing the simple technology needed to capture and destroy HFC-23. This technological solution would cost the developed world less than €100 million, saving an estimated €4.6 billion in CDM credits that could be spent on other climate-protecting uses. Similar technological fixes could work for industrial emissions of nitrous oxide from nylon feedstock and fertilizer manufacture.

Trading places

Supporters of HFC-23 projects argue that the entire point of the CDM is to identify low-cost opportunities to reduce emissions, and once identified, they should not be skimmed off the top of the market. But the CDM is both a market and a subsidy from industrialized to developing countries. As a subsidy, it should be judged by how effectively it reduces emissions for each dollar expended. In these terms, the CDM is a very inefficient subsidy. An alternative mechanism for reducing HFC-23 would require a separate protocol to the United Nations Framework Convention on Climate Change, but would be administratively tractable because of the small number of installations involved. Indeed, a similar mechanism has proven successful in compensating developing nations for the cost of switching from ozone-depleting substances under the Montreal Protocol.

Future emissions scenarios suggest that unless China and India can be convinced to build mostly efficient, low-carbon-emitting electricity-generating plants from natural gas rather than coal over the next one to two decades, little can be done to stem the tide of global climate change. Perversely, the presence of cheap non-CO_2 credits such as HFC-23 in the market is a disincentive to developing new carbon-limiting energy projects that would help to achieve this goal.

"The global carbon market is a very inefficient subsidy."

There is an obvious solution to what is wrong with the CDM: make the global carbon market a market for CO_2 rather than for all six Kyoto Protocol gases. The first two years of the CDM have generated high participation that could be harnessed to put the developing world, especially China and India, on a path to a low-carbon future. The existing structure of the carbon market is fixed for the period of the Kyoto Protocol. To attempt a change in mid-course would alarm investors, but European Union governments as well as Japan can send a clear signal that after 2012, they are interested in purchasing CO_2-only credits, and that preference will be given to projects in the energy sector.

Given sufficient warning, the energy sector in China and India will probably meet this new demand for low-cost carbon credits from the developed world. Industrial emissions of HFC-23, nitrous oxide and methane should, at the same time, be addressed by a separate agreement that fully compensates producers of these gases for the cost of abating emissions. Rich nations would save money by paying the actual cost of abatement rather than inflated market prices, and use these savings for further climate abatement through the CDM or other policies and measures.

But fixing the carbon market is unlikely to be enough to put major developing nations on a path to low-carbon energy. Because the CDM awards credits for the difference between baseline and actual emissions from a project, its impact will always be marginal. Ultimately, it is the baseline emissions path that must be altered if the problem of global warming is to be resolved.

What matters in the long term is the type of energy infrastructure that gets locked into place in the world economy. Tackling that problem requires identifying economic, national security, as well as energy priorities of the major developing economies and then finding ways to align them with low-carbon energy infrastructures[5]. The CDM, no matter what the price of carbon, is unlikely to convince China that it makes more sense to depend on foreign sources of natural gas than on cheaper domestic coal. Similarly, India is unlikely to pursue nuclear energy to significantly reduce its carbon emissions, given the challenges of non-proliferation and nuclear waste, without greater international support[6].

The CDM might have a role to play here by creating a secure market for future technology for low-carbon energy. But this won't happen while market resources are diverted into abating waste gases associated with the refrigerant, nylon and fertilizer industries. In the period beyond 2012, signatories to Kyoto should recognize that measures in addition to the global carbon market are needed to set the developing world on a path towards a sustainable-energy future. These include substantial increases in technology investment, agreements to share low-carbon technologies as they are developed and a commitment to fostering resilient energy markets and security arrangements so that it is in the interest of key developing nations to foster low-carbon economic growth. ■

Michael Wara was at the Program on Energy and Sustainable Development, Stanford University, Stanford, California 94305, USA, and is now at Holland & Knight LLP, 50 California Street, San Francisco, California 94111, USA.

1. DEFRA EU Emissions Trading Scheme, UK National Allocation Plan at www.defra.gov.uk/environment/climatechange/trading/eu/phaseII-nap.pdf (2006).
2. Victor, D. G., House, J. & Joy, S. *Science* **309,** 1820–1821 (2005).
3. Energy Information Administration *International Energy Outlook 2006* at www.eia.doe.gov/oiaf/ieo/index.html (2006).
4. Wara, M. *Measuring the Clean Development Mechanism's Performance and Potential* at http://pesd.stanford.edu/publications/cdm (2006).
5. Heller, T. C. & Shukla, P. R. *Development and Climate: Engaging Developing Countries* (Pew Center on Global Climate Change, Arlington, 2003).
6. Jackson, M., Joy, S., Heller, T. C. & Victor, D. *Greenhouse Gas Implications in Large Scale Infrastructure Investments in Developing Countries: Examples from India and China* at http://pesd.stanford.edu/publications/deals/ (2006).

See Business, page 584.

Lifting the taboo on adaptation

Renewed attention to policies for adapting to climate change cannot come too soon for **Roger Pielke, Jr**, **Gwyn Prins**, **Steve Rayner** and **Daniel Sarewitz**.

During the early policy discussions on climate change in the 1980s, adaptation was understood to be an important option for society. Yet for much of the past two decades the mere idea of adapting to climate change became problematic for those advocating emissions reductions, and was treated "with the same distaste that the religious right reserves for sex education in schools. That is, both constitute ethical compromises that in any case will only encourage dangerous experimentation with the undesired behaviour"[1]. Indeed, former US vice-president Al Gore forcefully declared his opposition to adaptation in 1992, explaining that it represented a "kind of laziness, an arrogant faith in our ability to react in time to save our skins".

But perspectives have changed. Adaptation is again seen as an essential part of climate policy alongside greenhouse-gas mitigation. Both the recent Stern Review on the Economics of Climate Change[2] and the efforts of the Intergovernmental Panel on Climate Change[3] demonstrate that adaptation is firmly back on the agenda. There are at least three reasons why the taboo on adaptation can no longer be enforced.

First, there is a timescale mismatch. Whatever actions ultimately lead to the decarbonization of the global energy system, it will be many decades before they have a discernible effect on the climate. Historical emissions dictate that climate change is unavoidable. And even the most optimistic emissions projections show global greenhouse-gas concentrations rising for the foreseeable future.

Second, vulnerability to climate-related impacts on society are increasing for reasons that have nothing to do with greenhouse-gas emissions, such as rapid population growth along coasts and in areas with limited water supplies. As Hurricane Katrina made devastatingly clear, climate vulnerability is caused by unsustainable patterns of development combined with socioeconomic inequity[4]. Post-Katrina debate focused on whether or not the event bore the signature of global warming, despite the fact that scientists have known for decades the inevitability of a Katrina-like disaster in New Orleans.

Finally, those who will suffer the brunt of climate impacts are now demanding that the international response to climate change focus on increasing the resilience of vulnerable societies to damaging climate events that — like Katrina — will occur regardless of efforts to mitigate emissions. In 2002, developing countries put forward the 'Delhi Declaration', calling for greater attention to adaptation in international climate-change policy negotiations[5].

R. RANOCO/REUTERS/CORBIS

Rising sea levels are only part of the story — in the Phillipines, land use changes have increased flood risk.

Mind the gap

The rehabilitation of the idea of adaptation is overdue and seems straightforward. But there is an elephant in the room: the core assumptions underlying contemporary climate-change policy conflict with the goal of increasing resilience to natural climate change and variability. Adaptation cannot just be dusted off and embraced — new ways of thinking about, talking about and acting on climate change are necessary if a changing society is to adapt to a changing climate.

> **"New ways of thinking about, talking about and acting on climate change are necessary if a changing society is to adapt to a changing climate."**

The United Nations (UN) Framework Convention on Climate Change (UNFCCC) treats adaptation in the narrowest sense — as actions taken in response to climate changes resulting from anthropogenic greenhouse-gas emissions[6]. By contrast, for decision-makers and researchers focused on sustainable development and disaster mitigation, adaptation describes a much broader range of actions that make societies more robust to changes, including, but not limited to, those caused by climate change[6].

This distinction profoundly affects society's ability to take effective action. In the UNFCCC's view, adaptation is only necessary because of greenhouse-gas emissions — an interpretation that is widely accepted. For instance, the Stern Review explained that adaptation "is crucial to deal with the unavoidable impacts of climate change to which the world is already committed"[2]. Adaptation, therefore, represents a cost of human-caused climate change that would be avoided if climate change were prevented through emissions mitigation.

At the margins

But most projected impacts of anthropogenic climate change are marginal increases on already huge losses. Locating adaptation in this margin creates bizarre distortions in public policy. For example, in the Philippines, policymakers have begun to acknowledge the flood threats posed by the gradual sea-level rise of 1 to 3 millimetres per year, projected to occur with climate change. At the same time, they remain oblivious to, or ignore, the main reason for increasing flood risk: excessive groundwater extraction, which is lowering the land


K. SENOSI/AP

Malaria risks are increasing for reasons that have nothing to do with climate change.

surface by several centimetres to more than a decimetre per year[7]. As with Katrina, the political obsession with the idea that climate risks can be reduced by cutting emissions distracts attention from the more important factors that drive flood risks.

Similarly, non-climate factors are by far the most important drivers of increased risk to tropical disease. For instance, one study found that without taking into account climate change, the global population at risk from malaria would increase by 100% by 2080, whereas the effect of climate change would increase the risk of malaria by at most 7% (ref. 8). Yet tropical disease risk is repeatedly invoked by climate-mitigation advocates as a key reason to curb emissions. In a world where political attention is limited, such distortions reinforce the current neglect of adaptation.

The wrong direction

Virtually every climate impact projected to result from increasing greenhouse-gas concentrations — from rising storm damage to declining biodiversity — already exists as a major concern. As long as adaptation is discussed in terms of its marginal effects on anthropogenic climate change, its real importance for society is obscured.

The focus on mitigation has created policy instruments that are biased against adaptation[6]. Under the Kyoto Protocol to the UNFCCC, rich countries pay costs that poor countries incur by adapting to the marginal impacts of climate change — but they can in principle avoid these costs through enhanced mitigation efforts[9]. This provision of the Protocol exemplifies the failure to take adaptation seriously: not only are the funds involved provided on a voluntary basis by rich countries but they are held hostage to mitigation[9]. The logic is that greenhouse-gas reductions will, in turn, reduce marginal adaptation costs. In practice, this means that the UNFCCC will pay "costs that lead to global environmental benefits, but not those that result in local benefits".[9] To those experiencing devastating losses from climate impacts in developing countries, such logic must sound surreal: policy 'success' means not investing in adaptation even as climate impacts, driven mainly by non-climate factors, continue to mount.

"Policy-makers need to understand the limitations of mitigation for reducing vulnerabilities."

To address the bias in the Kyoto Protocol, some have suggested that a new protocol, focused on adaptation, be developed under the Climate Convention[10], but it does not seem to have wide support. Others suggest that adaptation be 'mainstreamed' into existing institutions focused on sustainable development and disaster reduction, such as the UN International Strategy for Disaster Reduction[11]. The reality, of course, is that adaptation is already mainstreamed. The roof over your head, complex reinsurance contracts for disasters or, indeed, every other institution, technology and policy that helps people to live safely and prosperously in the face of climatic variability, change and uncertainty are mainstream. The challenge is to move more of humanity into this mainstream.

Taking responsibility

Progress on adaptation is also distorted by the common assumption that marginal adaptation is a local issue, whereas mitigation is a global one, requiring global coordination[2,9]. But does the distinction hold? With the ongoing failure of many rich nations to reduce emissions, action on mitigation has become increasingly diffuse as communities, cities, states and companies pursue emissions reductions. By contrast, the absence of a high-profile international vehicle for focusing attention on the broad benefits of adaptation seems to be one reason for its poor-cousin status at all scales of policy making[11].

What would a more vigorous international debate on adaptation bring? Those who have been concerned that attention to adaptation (and sustainable development) would detract from mitigation efforts have sought to avoid such a debate. Yet policy-makers need to understand the limitations of mitigation for reducing vulnerabilities, and give more urgent consideration to broader adaptation policies — such as improved management of coastal zones and water resources — that will enhance societal resilience to future climate impacts regardless of their cause. To define adaptation as a cost of failed mitigation is to expose millions of poor people in compromised ecosystems to the very dangers that climate policy seeks to avoid.

A poor fit

But defining adaptation in terms of sustainable development does not fit comfortably into the current political framework of the climate-change problem. By introducing sustainable development, one is forced to consider the missed opportunities of an international regime that for the past 15 years or more has focused enormous intellectual, political, diplomatic and fiscal resources on mitigation while downplaying adaptation. Until adaptation is institutionalized at a level of intensity and investment at least equal to those of the UNFCCC and Kyoto Protocol, climate impacts will continue to mount unabated, regardless of even the most effective cuts in greenhouse-gas emissions. ■

Roger Pielke, Jr, is at the University of Colorado. Gwyn Prins is at the London School of Economics and Columbia University. Steve Rayner is at Oxford University's James Martin Institute. Daniel Sarewitz is at Arizona State University.
e-mail: pielke@colorado.edu

1. Thompson, M. & Rayner, S. Cultural Discourses in *Human Choice and Climate Change* Vol. 1 (Battelle Press, Columbus, 1998).
2. Stern, N. (ed.) *The Economics of Climate Change: The Stern Review* (Cambridge Univ. Press, Cambridge, 2006).
3. Intergovernmental Panel on Climate Change http://www.ipcc.ch.
4. Pielke, R. A. Jr & Sarewitz, D. *Population and Environment* **26,** 255–268 (2005).
5. UNFCCC *The Delhi Declaration on Climate Change and Sustainable Development* at http://unfccc.int/cop8/latest/delhidecl_infprop.pdf (October 2002).
6. Pielke, R. A. Jr *Environmental Science & Policy* **8,** 548–561 (2005).
7. Rodolfo, K. S. & Siringan, F. P. *Disasters* **30,** 118–139 (2006).
8. Goklany, I. *Science* **306,** 55–57 (2004).
9. Bouwer, L. & Aerts, J. C. J. H. *Disasters* **30,** 49–63 (2006).
10. Burton, I. *Policy Options* December-January, 33–38 (2005).
11. Huq, S. & Reid, H. *IDS Bulletin* **35,** 15–21 (2004).

See Editorial, page 567.

BOOKS & ARTS

A man of magnitude

Charles Richter developed the scale for measuring earthquakes.

Richter's Scale: Measure of an Earthquake, Measure of a Man
by Susan Elizabeth Hough
Princeton University Press: 2007. 336 pp. $27.95, £17.95

Gregory C. Beroza

The cover of *Richter's Scale* by Susan Hough features an early photograph of Charles Richter looking like a grown-up Harry Potter. Like much of the rest of this biography, it reveals an unfamiliar side of the scientist famous for developing the first magnitude scale for earthquakes in 1935.

Until now, younger seismologists have known of Richter's life primarily by word of mouth. This has led to a rather uneven treatment. For example, I have been told many times that Richter was an avid nudist, but I was unaware that he was an alumnus of the Stanford University physics department. As Hough's biography repeatedly makes clear, there is much more to this intensely private and extremely complicated man.

Hough, a seismologist at the United States Geological Survey in Pasadena, California, brings a distinctly scientific approach to her writing, systematically presenting the data, learning what she can from it, and stating the uncertainties and non-uniqueness in her interpretations. Her data include Richter's personal and professional papers from the archives at the California Institute of Technology, his poetry (he was a dedicated writer of verse) and interviews with those who knew him.

The book chronicles the many difficulties in his family life, which at times are so tortuous that they are hard to follow. It also covers his many idiosyncrasies and indiscretions. In exploring the latter, Hough extrapolates somewhat to infer both Richter's intentions and his actions. This leads her onto shaky ground, and at times feels uncomfortably invasive, but she is careful to express both her reasoning and the uncertainty in her interpretations. She speculates that much of Richter's behaviour could be explained by Asperger's syndrome, a form of autism — a diagnosis that was not available during his lifetime.

The biography shines brightest in its treatment of scientific matters. It may come as some surprise that, among seismologists, there is a notion that Richter is more famous than he ought to be. Why and how do scientists become famous? In the case of seismology, earthquakes provide opportunities for scientists to speak to a wide audience. Richter clearly took advantage of this platform to promote an awareness of earthquake hazards. However, there can be a fine line between public outreach and self-promotion. *Richter's Scale* explores these issues, which transcend any single discipline, and sets the record straight in Richter's case.

CALTECH SEISMOLOGICAL LAB.

Charles Richter and the measurement of quakes: a portrait of the young seismologist in his 'Harry Potter' days.

Conventional wisdom holds that credit for developing the magnitude scale ought to rest equally with Richter and Beno Gutenberg, a colleague of Richter's at Caltech who is recognized as one of the giants of seismology. Hough persuasively argues that credit for developing the first earthquake magnitude scale rightly belongs to Richter. She also conveys a sense of what an undertaking it must have been to establish an earthquake magnitude scale given technological limitations of the 1930s. The development of subsequent magnitude scales that can be applied more generally, and their interpretation in terms of the energy released in earthquakes, was clearly a collaborative effort, with Gutenberg taking the lead role. The unfortunate practice of referring to all magnitude scales as 'the Richter scale' contributes to the perceived slight of Gutenberg.

Magnitude scales are not the be-all and end-all of seismology. What about Richter's wider contributions? Here too, Hough makes a case that he has been underappreciated. Richter was a full partner in the systematic studies of wave propagation and global earthquake activity that led, in addition to journal publications, to Gutenberg and Richter's seminal *Seismicity of the Earth and Associated Phenomena* (Princeton University Press, 1954). Richter also wrote his own landmark textbook, *Elementary Seismology* (Freeman, 1958). An understanding of the relative frequency of large and small earthquakes and the implications for earthquake risk are first expressed in Richter's original paper on magnitude in the *Seismological Society of America Bulletin* in 1935, but it was nearly a decade before he and Gutenberg quantified this observation and published their result.

Hough's book concludes with a draft of the speech Richter planned to give at his retirement. The speech serves as a metaphor for his hidden life, covering a wide range of topics in science, the arts, history and philosophy. Its seismological aspects are preoccupied with seismic risk reduction, an activity Richter championed in the latter part of his career. His comments are clearly, if somewhat bluntly, stated. He never delivered the speech, however, owing to his discomfort at the way his retirement party played out, so most of his colleagues remained unaware of the scope of his thoughts and interests. *Richter's Scale* will change this. It reveals Richter to be an individual with more than his share of flaws, but also as an iconoclastic scientist worthy of his fame and of our admiration. ■

Gregory C. Beroza is in the Department of Geophysics, Stanford University, Stanford, California 94305, USA.

Unfit for modern life

Mismatch: Why Our World No Longer Fits Our Bodies
by Peter Gluckman & Mark Hanson
Oxford University Press: 2006. 304 pp.
£16.95, $29.95

Michael Sargent

Will people born in the developed world in the 1990s have shorter lives than those born 60 years ago? If the fears of the authors of the thought-provoking book *Mismatch* are realized, the upward trend in life expectancy over the past century is set to reverse unless the lifestyle of young people dramatically improves. The authors, Peter Gluckman and Mark Hanson, are not alone in their anxiety: some US experts on longevity believe that the steady rise of childhood obesity in their country is likely to reduce life expectancy by 2 to 5 years by the middle of the century, with enormous increases in morbidity from type 2 diabetes, heart disease and other vascular disorders.

Curiously, while the horror of this impending public-health disaster lies in the future, the authors overlook the tragedy already afflicting the Pima Indians of Arizona and the inhabitants of the Pacific island of Nauru. Just a few decades ago, these populations were lean and fit and knew nothing of diabetes; now the majority of young adults become obese, develop diabetes and die prematurely of heart disease. A genetic predisposition to diabetes may have underpinned the crisis, but the immediate cause was undoubtedly an abrupt decline in strenuous physical activity along with the displacement of the traditional cuisine by energy-dense food.

Recognizing the emergence of a global epidemic of a similar kind, Gluckman and Hanson condemn the mismatch between the intrinsic physiological capacities programmed *in utero* and our twenty-first-century lifestyles, which encourage the consumption of excess calories without the physical demands of former times. With Old Testament severity the authors call for "a return to a different way of life" and condemn the modern habitat, to which they feel we are poorly adapted — a counterintuitive notion at a time when life expectancy is higher than ever.

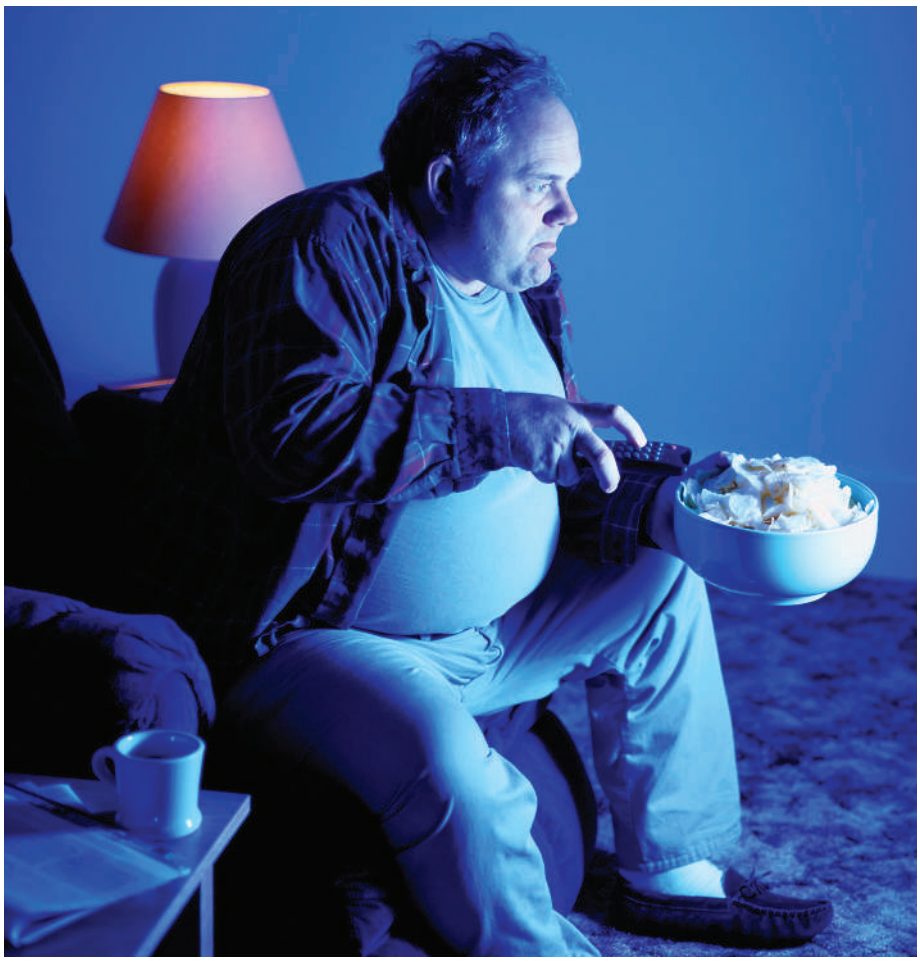
The heart of the book is predicated on David Barker's idea that nutritional and other kinds of stress impinging on pregnant women may register *in utero* and affect the postnatal development of progeny. Extensive investigations of laboratory mammals and human epidemiology portray the response to nutritional stress as a kind of faustian bargain. The cost of protecting the brain and reproductive behaviour is underdeveloped viscera that may compromise health and survival in later life. There is a greater burden if the nutritional experiences of the offspring are richer than their mother's during pregnancy, with greater risks of premature onset of the chronic diseases of middle age: coronary heart disease, type 2 diabetes, hypertension and osteoporosis. The embryo is also sensitive to excessive nutrition from high maternal blood sugar and psychological stress, resulting in epigenetic modifications of genomic function that are expressed after birth in surprising ways in humans and animals.

The average stature and physique of humans from different parts of the world reflect the local nutritional regime by processes that are established *in utero* and transmitted epigenetically to succeeding generations. It seems that obesity and its consequences are likely to emerge when this regime changes. In the developing world, those released from the grip of historic poverty are most affected, whereas in the developed world, for reasons that are less well understood, the victims are the least educated and the least affluent. The authors believe that research should focus on improving maternal nutrition to ensure that babies are better adapted to changing nutritional landscapes.

Gluckman and Hanson also apply their 'mismatch paradigm' to a diverse group of situations in which cultural developments that challenge the logic of darwinian evolution generate new problems. These include the emergence of a long post-reproductive life, and the tricky interval between sexual maturity and the time when young adults are considered mature enough, for example, to hire a car. Other issues include attitudes to breastfeeding, the effect of breeding patterns on cancers of the reproductive organs, living with micronutrient deficiencies, and the onset of myopia provoked by excessive reading at an early age. Other authors mining the same vein have unearthed more issues of human caprice and mismatches of the kind the authors find so disturbing.

The authors sometimes skim over human biology, preferring to explore the natural history of other species. I was amused by their conflation of Psalm 90 with the Gettysburg address, and horrified by persistent references to the 'design' of organisms — a usage notably obstructive to an understanding of the evolutionary process, the disclaimer notwithstanding. Overall, however, this book conveys admirably, for a non-specialist reader, the implications of an important idea. ■

Michael Sargent is a developmental biologist at the National Institute for Medical Research, Mill Hill, London NW7 1AA, UK. He is the author of *Biomedicine and the Human Condition: Challenges, Risks, and Rewards*.



The fat of the land: obesity stems from our failure to adapt to an increasingly high-calorie diet.

A big bite of the past

By standing up for themselves between 3 million and 4 million years ago, Lucy and her fellow *Australopithecus afarensis* caused quite a stir. But bipedalism is just one factor in the rich mix of human evolution, as amply shown in the revised, updated and expanded *From Lucy to Language* (Simon & Schuster, $65).

Donald Johanson, who discovered Lucy, and his co-writer Blake Edgar have added the big finds since 1996 to their brilliant overview, including the Indonesian 'hobbit' *Homo floresiensis*. And as this snap of *A. afarensis* teeth from Ethiopia reveals, the expanded range of photos — many at actual size — remain jaw-droppingly spectacular. **B.K.**

J. READER/SPL

Back to basics

Darwinian Reductionism: Or, How to Stop Worrying and Love Molecular Biology
by Alex Rosenberg
University of Chicago Press: 2006. 272 pp. $40, £25.50

Bruce H. Weber

The understanding we have gained about the molecular basis of living systems and their processes was a triumph of twentieth-century science. Since the structure of DNA was elucidated in 1953, molecular biologists have been deepening our insights into a wide range of biological phenomena. It has been a heady time: it seemed that mendelian genetics would be reduced to the macromolecular chemistry of nucleic acids, with biology set to become a mature science in the same way as physics and chemistry. The emerging field of the philosophy of biology inherited the reductionist framework of logical empiricism. But as our knowledge of molecular biology deepened, many philosophers of biology, including David Hull, Philip Kitcher, Eliot Sober, Evelyn Fox-Keller and Paul Griffiths, saw that the reductionist approach faced serious problems.

There is no simple correlation between the mendelian gene and the increasingly complex picture provided by molecular genetics. To make matters worse, the theory to be reduced was presumably the population-genetic version of darwinian natural selection, which had from the start excluded phenomena about development and their possible link to evolutionary dynamics. Given this absence, Ernst Mayr, a founder of the modern evolutionary synthesis, argued that, although biological systems did not violate the laws of chemistry and physics, evolving biological systems have properties that cannot be reduced to such laws. The crux of the issue as Mayr saw it was that, whereas the physical sciences deal only in proximate explanations, the biological sciences also deal with ultimate explanations relating to evolutionary descent and the action of selection to produce adapted function. This, Mayr argued, resulted in the autonomy of biology with respect to the physical sciences. Alex Rosenberg's book *Darwinian Reductionism* is a response to the anti-reductionist position in contemporary philosophy of biology and to the autonomist stance of some biologists.

Rosenberg's thesis is that biological phenomena, including their functional aspects, are best understood at the level of their macromolecular constituents and their interactions in cellular environments that are themselves made up of other molecules. This has been, and continues to be, he argues, a successful, progressive research programme. He focuses in particular on the great advances in our understanding of developmental molecular biology, which teaches us how the genes that are involved in development function, interact and work with chemical gradients, for example, to produce morphology. Rosenberg provides an accessible review of current ideas on the 'wiring' of such gene complexes and the way they help account for morphological evolution. He is one of the first philosophers to consider the implications of 'evo-devo' (evolutionary developmental biology), and seizes the opportunity to promote a reductionist interpretation that was simply not possible with population genetics.

He shows a good grasp of the scientific details of developmental molecular biology, but it is unfortunate that in the introduction he gets the molecular details of sickle-cell anaemia wrong and then describes a resulting arterial blockage, rather than the lysis of red blood cells. This should not have survived the reviewing and editing process, but it is the only serious lapse. When he returns to the issue of mutant haemoglobins later in the book, he gets the molecular details for sickle-cell haemoglobin correct.

To bridge Mayr's gap between ultimate (natural selection) causes and proximate (structural and functional) causes, Rosenberg cites Theodosius Dobzhansky's dictum that nothing in biology makes sense except in the light of evolution. The various molecules in cells and the gene sequences of the macromolecules are products of previous selection by which their proximately causal (structural and functional) properties were screened. In bringing causality to bear on explanation, he makes use of the distinction between 'how possible' and 'why necessary' explanations. Ultimate historical explanations of current biological structures and functions are 'how possible' in type. But why particular molecular arrangements were selected in the past has the force of 'why necessary' explanation. This removes the burden from selectional dynamics of having to be predictive in order to be reductionist.

Rosenberg realizes that theory reductionism requires the theory of darwinian natural selection to be grounded in, or reduced to, a principle of natural selection at the level of chemical systems in which both stability and replicability are selected for. In effect, he produces a scenario in which biological selection can be reduced to chemical selection during the origin of life. This crucial move needs more careful analysis than Rosenberg provides. He gives, in effect, a 'how possible' explanation for the emergence of life and biological selection, but not a 'why necessary' one. For that he would need to deal with the literature of the origin of life and the more general recent work on complexity. Such an investigation would show that phenomena in these areas are more emergent than Rosenberg believes, and that there is a need to develop a theory of organization and emergence. Research on emergent complexity is still a work in progress, but it may undercut Rosenberg's thesis by providing a fully naturalistic, non-reductionist account of emergence. Such a non-reductionist account would not be anti-reductionist in the sense Rosenberg uses the term, but would offer a 'why necessary' explanation of the emergent phenomena. ■

Bruce H. Weber is emeritus professor in the Department of Chemistry and Biochemistry, California State University, Fullerton, and in the Division of Science and Natural Philosophy, Bennington College, Bennington, Vermont, USA.

A clash of two cultures

Physicists come from a tradition of looking for all-encompassing laws, but is this the best approach to use when probing complex biological systems?

Evelyn Fox Keller

Biologists often pay little attention to debates in the philosophy of science. But one question that has concerned philosophers is rapidly coming to have direct relevance to researchers in the life sciences: are there laws of biology? That is, does biology have laws of its own that are universally applicable? Or are the physical sciences the exclusive domain of such laws?

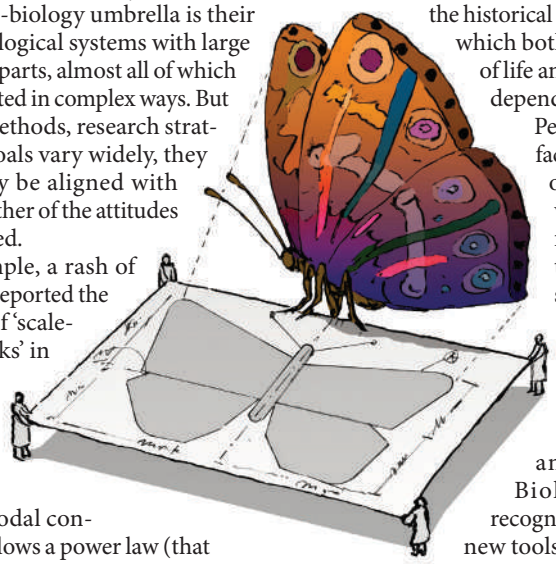
Today, biologists are faced with an avalanche of data, made available by the successes of genomics and by the development of instruments that track biological processes in unprecedented detail. To unpack how proteins, genes and metabolites operate as components of complex networks, modelling and other quantitative tools that are well established in the physical sciences — as well as the involvement of physical scientists — are fast becoming an essential part of biological practice. Accordingly, questions about just how much specificity needs to be included in these models, about where simplifying assumptions is appropriate, and about when (if ever) the search for laws of biology is useful, have acquired pragmatic importance — even some urgency.

In the past, biologists have been little concerned about whether their findings might achieve the status of a law. And even when findings seem to be so general as to warrant thinking of them as a law, the discovery of limits to their generality has not been seen as a problem. Think, for example, of Mendel's laws, the central dogma or even the 'law' of natural selection. Exceptions to these presumed laws are no cause for alarm; nor do they send biologists back to the drawing board in search of better, exception-free laws. They are simply reminders of how complex biology is in reality.

Physical scientists, however, come from a different tradition — one in which the search for universal laws has taken high priority. Indeed, the success of physics has led many to conclude that such laws are the *sine qua non* of a proper science, and provide the meaning of what a 'fundamental explanation' is.

Physicists' and biologists' different attitudes towards the general and the particular have coexisted for at least a century in the time-honoured fashion of species dividing their turf. But today, with the eager recruitment of physicists, mathematicians, computer scientists and engineers to the life sciences, and the plethora of institutes, departments and centres that have recently sprung up under the name of 'systems biology', such tensions have come to the fore.

Perhaps the only common denominator joining the efforts currently included under the systems-biology umbrella is their subject: biological systems with large numbers of parts, almost all of which are interrelated in complex ways. But although methods, research strategies and goals vary widely, they can roughly be aligned with one or the other of the attitudes I've described.



J. KAPUSTA/IMAGES.COM

For example, a rash of studies has reported the generality of 'scale-free networks' in biological systems. In such networks, the distribution of nodal connections follows a power law (that is, the frequency of nodes with connectivity k falls off as $k^{-\alpha}$, where α is a constant); furthermore, the network architecture is assumed to be generated by 'growth and preferential attachment' (as new connections form, they attach to a node with a probability proportional to the existing number of connections). The scale-free model has been claimed to apply to complex systems of all sorts, including metabolic and protein-interaction networks. Indeed, some authors have suggested that scale-free networks are a 'universal architecture' and 'one of the very few universal mathematical laws of life'.

But such claims are problematic on two counts: first, power laws, although common, are not as ubiquitous as was thought; second, and far more importantly, the presence of such distributions tells us nothing about the mechanisms that give rise to them. 'Growth and preferential attachment' is only one of many ways of generating such distributions, and seems to be characterized by a performance so poor as to make it a very unlikely product of evolution.

How appropriate is it to look for all-encompassing laws to describe the properties of biological systems? By its very nature, life is both contingent and particular, each organism the product of eons of tinkering, of building on what had accumulated over the course of a particular evolutionary trajectory. Of course, the laws of physics and chemistry are crucial. But, beyond such laws, biological generalizations (with the possible exception of natural selection) may need to be provisional because of evolution, and because of the historical contingencies on which both the emergence of life and its elaboration depended.

Perhaps it is time to face the issues head on, and ask just when it is useful to simplify, to generalize, to search for unifying principles, and when it is not. There is also a question of appropriate analytical tools. Biologists clearly recognize their need for new tools; ought physical scientists entering systems biology consider that they too might need different methods of analysis — tools better suited to the importance of specificity in biological processes? Finally, to what extent will physicists' focus on biology demand a shift in epistemological goals, even the abandonment of their traditional holy grail of universal 'laws'? These are hard questions, but they may be crucial to the forging of productive research strategies in systems biology. Even though we cannot expect to find any laws governing the search for generalities in biology, some rough, pragmatic guidelines could be very useful indeed. ■

Evelyn Fox Keller is at the Massachusetts Institute of Technology, 77 Mass Avenue, E51-185, Cambridge, Massachusetts 02139, USA, and a Blaise Pascal chair in Paris, France.

FURTHER READING
Barabási, A. L. & Bonabeau, E. *Sci. Am.* **288,** 50–59 (2003).
Beatty, J. in *Concepts, Theories and Rationality in the Biological Sciences* (eds Lennox, J. G. & Wolters, G.) 45–81 (Univ. Pittsburgh Press, Pittsburgh, 1995).
Keller, E. F. *BioEssays* **27,** 1060–1068 (2005).
Keller, E. F. *Making Sense of Life: Explaining Biological Development with Models, Metaphors, and Machines* (Harvard Univ. Press, Cambridge, MA, 2002).

For other essays in this series, see http://nature.com/nature/focus/arts/connections/index.html

NEWS & VIEWS

QUANTUM PHYSICS

Indistinguishable from afar

Michael Fleischhauer

Imprinting a coherent light pulse on the spins of atoms is standard quantum sorcery. Retrieving the same light pulse from a second, distant set of atoms looks rather like black magic. But it, too, is just quantum mechanics.

In the quantum world, particles of the same kind are indistinguishable: the wavefunction that describes them is a superposition of every single particle of that kind occupying every allowed state. Strictly speaking, this means that we can't talk, for instance, about an electron on Earth without mentioning all the electrons on the Moon in the same breath. Fortunately, the existence of distant particles is for practical purposes inconsequential, and we can generally get away with ignoring them. On page 623 of this issue[1], however, Ginsberg, Garner and Hau show that two ensembles of atoms, known as Bose–Einstein condensates, can be made to reveal their indistinguishability, even when they are some distance apart. The condensates are admittedly not as far apart as Earth and the Moon, being separated by a fraction of a millimetre. But in the quantum world, that is a very large distance indeed.

To perform this trick, Ginsberg *et al.* employ a technique that was developed some years ago to 'store' a light pulse in an ensemble of atoms[2,3]. In the classical understanding of this process, an incoming laser pulse transfers its energy to a gas by inducing tiny, oscillating distributions of positive and negative electric charge — optical 'dipoles' — in the atoms. These dipoles radiate and quickly decay, so another laser is used to transfer the charge oscillations to oscillations in spin, which are more stable. When this control laser pulse is switched off, although the probe pulse disappears, its coherent information content is conserved in the spin oscillations of the atoms. If the control laser is then switched back on, this process is reversed[4]: the atoms radiate coherently, rather like an array of phased antennas, according to the phase of the original pulse imprinted in their spin oscillations. Light leaves the ensemble again as a coherent pulse with a well-defined shape and direction of propagation.

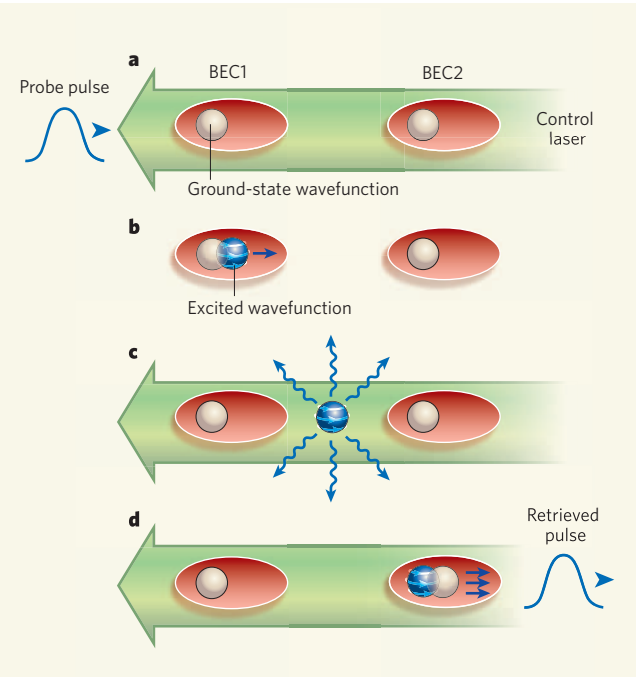


Figure 1 | Distant bounds. a, For their demonstration of quantum-mechanical indistinguishability on a macroscopic scale[1], Ginsberg, Garner and Hau use two ensembles of atoms known as Bose–Einstein condensates (BECs). With the help of a control laser, they store a probe-laser pulse in the first BEC with its atoms in their ground state. **b**, When the control laser is switched off, the probe-laser pulse imprints itself on the atoms of the condensate by exciting oscillating spin dipoles. The quantum-mechanical wavefunction of these dipoles consists of two superposed components corresponding to atoms in the ground and excited states. Momentum conservation requires that the excited-state wavefunction moves off as the excited atoms absorb their photons, but the ground-state wavefunction stays where it is. **c**, The initial pulse cannot be retrieved coherently while the spin-excited wavefunction is in transit between the two condensates; if the control laser is switched back on, only spontaneous, incoherent emission of photons occurs in all directions. **d**, If, on the other hand, the control laser is not switched back on until the excited-state wavefunction is within the second condensate, the original light pulse can be regenerated. The condensates were prepared independently so they would appear alien to each other. But they are composed of the same type of atoms in their ground state, and these are quantum-mechanically indistinguishable objects.

Quantum mechanically, spin dipoles are a superposition of the ground state and a spin-excited state of the atoms. The simultaneous existence of both components allows phase information to be stored, and is thus crucial to the coherence of the retrieval process. If only the spin-excited state were occupied, the atomic dipoles would oscillate with random phases, and when the control laser was switched on again it would induce only the spontaneous emission of incoherent light.

It is important to recall here that it is only atoms in the spin-excited state that have actually absorbed a photon. This is significant because, although the atoms of a Bose–Einstein condensate are very cold, they can move freely, and the process of storing the light pulse in the atomic ensemble is accompanied by a transfer of photon momentum to the excited atoms. As a result, the spin-excited part of the wavefunction starts to move away from its original position with the recoil velocity of the photon, whereas the ground-state part stays where it is. In a condensate, spin oscillations live long enough for the spin-excited component to travel over distances much larger than its own spatial extent. This leads to the question: might a coherent retrieval of the initial light pulse be possible even when the two components of the wavefunction — spin-excited and ground state — have become separated in space?

Ginsberg and colleagues[1] show that this can indeed be the case — at least as long as the wavepacket of the spin-excited states does not leave the condensate. This is because, in a Bose–Einstein condensate, the wavefunction of the ground state extends over the whole condensate. There is thus always a spatial overlap between it and the spin-excited wavefunction, and the retrieval process is always coherent.

But it is in a second experiment that the authors made their most astonishing observation. They prepared two independent

Bose–Einstein condensates in two traps, each extending over about 50 μm and separated by a fraction of a millimetre. A coherent light pulse was stored in the first ensemble, in the conventional way (Fig. 1a). A wavepacket of spin-excited atoms was created, which travelled, as before, with photon recoil velocity in the direction of the second condensate (Fig. 1b). Once this 'messenger' wavepacket had left the first condensate, a coherent regeneration of the original light pulse was no longer possible (Fig. 1c) — as expected.

But a strange thing happened if the time of retrieval was chosen to be much later — late enough for the messenger wavepacket to have reached the second condensate. Under these circumstances, the initial light pulse could indeed be regenerated, with well-defined shape and direction of propagation (Fig. 1d). The two condensates had been independently prepared, and so from a naive point of view the messenger wavepacket transferred from the first condensate should be completely alien to the atoms of the second. The only way to understand the experimental observation is to consider the atoms in both condensates as indistinguishable quantum objects. As such, the ground-state wavefunction would have a component in both traps simultaneously, and could thus combine with the spin-excited messenger component, once this had passed between the condensates, to coherently retrieve from the second condensate the light pulse stored in the first.

The work of Ginsberg *et al.*[1] is a striking and intriguing demonstration of a fundamental aspect of quantum physics — indistinguishability. But it also shows that we are entering a state of unprecedented experimental control of coherent light and matter waves. That could bring very real technological benefits: applications that spring to mind include quantum-information interfaces that allow the transfer of a quantum bit encoded in a photon to a single atom, as well as ultra-sensitive rotation sensors and gravity detectors. ■

Michael Fleischhauer is in the Fachbereich Physik, Technische Universität Kaiserslautern, Erwin-Schrödinger-Straße, D-67663 Kaiserslautern, Germany.
e-mail: mfleisch@physik.uni-kl.de

1. Ginsberg, N. S., Garner, S. R. & Hau, L. V. *Nature* **445,** 623–626 (2007).
2. Liu, C., Dutton, Z., Behroozi, C. H. & Hau, L. V. *Nature* **409,** 490–493 (2001).
3. Phillips, D. F., Fleischhauer, A., Mair, A., Walsworth, R. L. & Lukin, M. D. *Phys. Rev. Lett.* **86,** 783–786 (2001).
4. Fleischhauer, M. & Lukin, M. D. *Phys. Rev. Lett.* **84,** 5094–5097 (2000).

CANCER BIOLOGY

Gone but not forgotten

Norman E. Sharpless and Ronald A. DePinho

The p53 tumour-suppressor protein is a cell's principal guardian against cancer. Most cancers eliminate p53 — but it seems that its pathway remains intact, so resurrecting it might provide a cancer therapy.

Once fully established, many cancers rely for their maintenance on the persistent activation of certain cancer-promoting genes (oncogenes)[1,2]. Drug developers hope to exploit this dependency, as it makes tumours vulnerable to inhibitors of the associated oncogenic proteins. Indeed, several such inhibitors have already proven to be effective cancer treatments. Three papers[3–5] now confirm an analogous idea for a tumour-suppressor protein that limits cancer growth*. The work reveals that persistent inactivation of the p53 tumour-suppressor pathway is similarly required for tumour maintenance, and thus opens new therapeutic avenues against cancer.

Cancer is the outcome of many mutations, each endowing tumour cells with more of the biological requirements for the full-blown disease. With so many genetic aberrations present in each cancer, it is difficult to distinguish the primary 'driver' events that trigger the cancer from secondary 'passenger' alterations. Moreover, even if the driver mutations are known, understanding how these lesions interact to induce cancer is a formidable task. Which mutations are crucial to the maintenance of an established malignancy, for instance? And which are required only during the initial genesis and progression of the cancer? Identification of the key 'targets' required for tumour maintenance has been a major challenge in drug development[6].

The work by Ventura *et al.*[3] and Xue *et al.*[4], on pages 661 and 656 of this issue respectively, and by Martins *et al.*[5], published in *Cell*, concentrates on the p53 protein. This tumour suppressor is one of the heavyweights of cancer biology, as direct inactivation of the gene that encodes it is the most common mutation in human cancer, and the pathway that it controls is probably compromised to some degree in all human cancers. This pathway normally helps the cell to respond to DNA damage, such as that caused by radiation or carcinogenic chemicals. DNA damage activates p53, which in turn induces the expression of proteins that halt the cell-division cycle to allow for repair. Activation of p53 can also initiate programmes of cell death (apoptosis) or permanent growth arrest (senescence) if the DNA damage is persistent and severe.

The latest studies describe three distinct genetic approaches to producing mice that lack p53 function, either by inactivating the gene encoding p53, or by interfering with production of the protein. Mice lacking p53 are highly prone to spontaneous and carcinogen-induced tumours. The new twist here is that these mice are engineered so that the dormant p53 gene can be reawakened in tumours by treating the animals with a particular chemical. Importantly, despite the different technical approaches and tumour types in the three studies, the reinstatement of p53 expression led universally to a prompt and impressive regression of established, *in situ* tumours.

The mere synthesis of p53 in a cell is not in itself sufficient to suppress tumours — the protein must also be stabilized and switched on, something that does not generally occur in normal cells. All three studies[3–5] found that restoring expression of p53 caused tumour regression. This indicates that there is some feature of a cancer that is sufficient to activate p53, and that this reinstatement of physiological p53 function is enough to halt tumour growth. Although the re-establishment of p53 expression in tumours may be even more effective when combined with stimuli that stabilize and/or activate p53 function (such as radiotherapy), these papers show that, even without such stimuli, tumour cells harbour the signals that can trigger the destructive power of p53 if it is present. We believe that this general conclusion — that established tumours remain persistently vulnerable to p53 tumour-suppressor function — is the most significant finding of the work.

How p53 carries out its anticancer function seems to differ according to the tumour type and its context. For example, reinstating p53 function in p53-deficient lymphomas (blood cancers) rapidly induces apoptosis[3,5]. By contrast, p53 reactivation in two types of solid tumour (soft tissue sarcoma and hepatocellular carcinoma) induces a potent growth arrest featuring hallmarks of cellular senescence[3,4]. Senescence is involved in suppressing the early steps of cancer development in several tissues (reviewed in ref. 7), and Ventura *et al.*[3] and Xue *et al.*[4] establish that ongoing resistance to senescence is also needed to maintain established tumours. It is not clear which features of a cancer determine whether its response to p53 activation is apoptosis or senescence. But both outcomes are associated with tumour regression and so could be of therapeutic benefit.

The tumour regression seen in the sarcomas and hepatocellular carcinomas was associated with senescence without apoptosis — but if there is no cell death, how does the tumour get smaller? Xue *et al.*[4] report an unexpected cause

*This article and the *Nature* papers concerned[3,4] were published online on 24 January 2007.

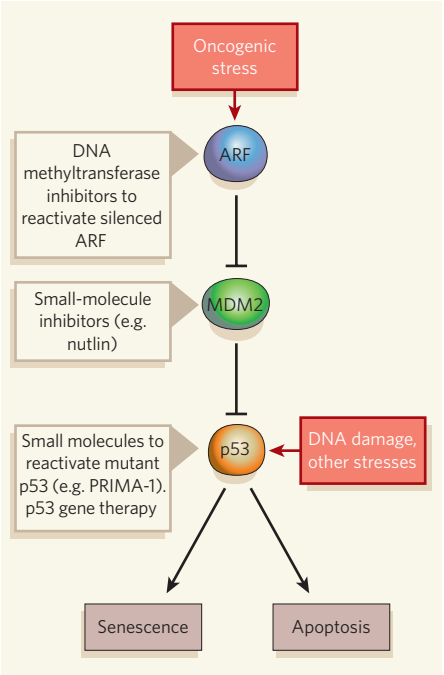


Figure 1 | Approaches to reactivating the p53 pathway. In the cell, p53 is usually degraded by the MDM2 protein. Various cellular stresses, including DNA damage, can activate p53, and once triggered, p53 can initiate programmed cell death (apoptosis) or permanent growth arrest (senescence). The p53 pathway is probably compromised to various extents in all human cancers. There are several potential approaches to reactivating the pathway in human cancers, depending on the type of lesion in the p53 pathway. Mutant p53 could be targeted directly with drugs or gene therapy. Or p53 signalling could be reactivated in tumours that overexpress MDM2 by using MDM2 inhibitors such as the nutlin molecules[12] that are in preclinical development. Likewise, the ARF tumour suppressor, which inhibits MDM2, is sometimes switched off in cancer by 'epigenetic' silencing, and agents that reverse such silencing are already well into human testing (for example, DNA methyltransferase inhibitors).

of tumour shrinkage: a rapid, non-apoptotic clearance of senescent tumour cells by an exuberant immune-mediated mechanism. As this tumour clearance takes place in 'nude' mice, which lack functional B and T immune cells, the authors argue that it represents activation of an innate immune response as a result of the production of proinflammatory molecules (for example, CSF1 and IL-15) by the senescent cells. It is worth noting, however, that markers of senescence accumulate in many tissues with age[8,9] and that senescent cells in precancerous tissue can persist for decades or more in humans[10,11]. Thus, some senescent cells seem to be impervious to this clearance mechanism. A better understanding of how this phenomenon works may make it possible to unleash the process on cancerous cells, while sparing normal ageing tissues.

These three papers[3–5] provide reason for cautious optimism that reactivation of p53, and possibly of other tumour-suppressor genes, might be useful in treating certain cancers. Reinstating p53 will no doubt be difficult in practice, but may be simplest in tumours with normal p53 that lack tumour-suppressor activity because of other mutations in the p53 pathway (Fig. 1). Moreover, the more rigorous determination of which factors are involved in tumour maintenance and the precise genetic and biological context in which they work will pave the way for further therapeutic possibilities.

The unwelcome finding in human patients, however, has been that although therapies that target tumour-maintaining oncogenes are initially effective, secondary genetic events occur all too often that render the tumours resistant to such treatment. Indeed, Martins *et al.*[5] inject a dose of clinical realism into the possibility of exploiting p53 reactivation for therapeutic ends. They found that although reactivation of p53 did induce widespread tumour-cell apoptosis, this was followed by the rapid appearance of tumours that progressed despite p53 expression. In many cases, these tumours harboured secondary lesions in regulators of the p53 pathway that rendered them resistant to p53. So, even under experimental conditions designed for near-optimal p53 reactivation, secondary resistance can limit the long-term benefit of the approach. Despite this cautionary observation, the three papers establish that resurrecting p53 can be of therapeutic benefit even in established, fully formed cancers. ■

Norman E. Sharpless is in the Department of Medicine and Genetics, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina 27599, USA.
Ronald A. DePinho is in the Center for Applied Cancer Science and Departments of Medical Oncology, Medicine and Genetics, Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115, USA.
e-mails: nes@med.unc.edu; ron_depinho@dfci.harvard.edu

1. Felsher, D. W. & Bishop, J. M. *Mol. Cell* **4,** 199–207 (1999).
2. Chin, L. *et al. Nature* **400,** 468–472 (1999).
3. Ventura, A. *et al. Nature* **445,** 661–665 (2007).
4. Xue, W. *et al. Nature* **445,** 656–660 (2007).
5. Martins, C. P., Brown-Swigart, L. & Evan, G. I. *Cell* **127,** 1323–1334 (2006).
6. Kola, I. & Landis, J. *Nature Rev. Drug Discov.* **3,** 711–715 (2004).
7. Narita, M. & Lowe, S. W. *Nature Med.* **11,** 920–922 (2005).
8. Dimri, G. P. *et al. Proc. Natl Acad. Sci. USA* **92,** 9363–9367 (1995).
9. Herbig, U., Ferreira, M., Condel, L., Carey, D. & Sedivy, J. M. *Science* **311,** 1257 (2006).
10. Michaloglou, C. *et al. Nature* **436,** 720–724 (2005).
11. Courtois-Cox, S. *et al. Cancer Cell* **10,** 459–472 (2006).
12. Vassilev, L. T. *et al. Science* **303,** 844–848 (2004).

PALAEOCLIMATE

When the world turned cold

Gabriel J. Bowen

As massive ice sheets grew on Antarctica during the first major glaciation of the Cenozoic era, the northern continents cooled and dried. The coincidence in timing implies that the cause was global rather than regional.

It has been nearly 34 million years since Earth was last free of large continental ice sheets. Before the beginning of the Oligocene epoch, 33.7 million years ago, Antarctica had been a lush, green continent for several tens of millions of years, and the only significant concentrations of ice on Earth probably occurred on the Antarctic highlands and in and around the Arctic Ocean. Then, in two short pulses spanning the first 300,000 years of the Oligocene, ice sheets grew over most of Antarctica[1]. The continent has been largely ice-covered ever since (Fig. 1, overleaf). But did the severe event at high southern latitudes affect the rest of the globe? And if so, how? Elsewhere in this issue, Dupont-Nivet *et al.*[2] and Zanazzi *et al.*[3] describe climatic and biotic changes in central Asia and North America during the early Oligocene that implicate declining levels of atmospheric carbon dioxide as a driver for changes in both hemispheres.

Given the importance of the early Oligocene event to the evolution of polar climate, it may be surprising that we have an imperfect understanding of its global consequences. In fact, records preserved in the veneer of sedimentary rock that coats the continents suggest a variety of changes in the climate, and in plant and animal communities, during the transition from the warm Eocene world to the cool Oligocene. But owing to inaccuracies in the dating of continental rocks and conflicting results generated using different methodologies, the interpretation of these records has been in question. Moreover, evidence for climate change in the Northern Hemisphere seemed to clash with the prevailing explanation for the cause of Antarctic glaciation in the early Oligocene: the tectonic rifting of Australia and South America from Antarctica, which was argued to have initiated the Antarctic Circumpolar Current and thus climatically isolated Antarctica[4]. Computer models suggest that such a change in ocean circulation would warm, rather than cool, the northern continents[5].

Dupont-Nivet *et al.*[2] (page 635) apply precision dating techniques to demonstrate that climate change in central Asia occurred at the

same time as the early Oligocene glaciation of Antarctica. Their data come from the Xining basin on the northeastern margin of the Tibetan plateau, where a thick stack of sedimentary rocks records the abrupt disappearance of lakes that had episodically flooded the basin prior to the Oligocene epoch. Using a detailed record of changes in Earth's magnetic field as these rocks were deposited, the authors temporally link the rock layers in the Xining basin to seafloor records documenting ice build-up in the Antarctic.

The magnetic records also demonstrate that the flooding and drying of the basin before the Oligocene drying event was episodic, and occurred in step with the same 100,000-year cyclic change in the shape of Earth's orbit that has been the pacemaker of recent ice ages. Counting these flood cycles gives the authors a precise assessment of the age of Xining basin drying relative to the marine records. The increase in Asian aridity was essentially synchronous with Antarctic glaciation. Furthermore, a major turnover of mammals is associated with the aridification of central Asia at this time[6], and Dupont-Nivet and colleagues' chronology provides the first solid link between this change and the global early Oligocene event.

In the second paper, Zanazzi *et al.*[3] (page 639) report new North American climate-change records produced using an innovative twist on a common geochemical technique. They measure the relative abundance of the stable isotopes of oxygen ($^{18}O/^{16}O$) in ancient minerals. This is determined by two factors: the $^{18}O/^{16}O$ of the water from which the mineral was derived, and the temperature at the time of formation. If one of these two values is known, the other can be estimated. The challenge for workers studying continental palaeoclimate is that usually neither parameter is known.

Zanazzi *et al.* overcome this limitation by analysing two materials that have different chemical histories. They use $^{18}O/^{16}O$ measurements of dense, chemically resilient tooth enamel from mammal fossils to estimate the $^{18}O/^{16}O$ of ancient water at their study sites in the northern Great Plains of the United States. Because the tooth mineral grew at the near-constant temperature of the mammal's body and — given its density — is unlikely to change in isotopic composition after deposition, the authors can estimate the $^{18}O/^{16}O$ of the animal's body water, which is closely related to that of its drinking water. Once the isotopic composition of water is known, they use $^{18}O/^{16}O$ from fossil bone, which is porous and obtains most of its oxygen at ambient temperatures after it is buried in soils, to reconstruct temperatures.

Zanazzi and colleagues' data[3] present a mixed picture relative to previous studies. They provide robust evidence for a substantial drop in average annual temperatures of about 8 °C. This is near the high end of values obtained in previous studies, and implies that the authors' continental site cooled much more than did the mid-latitude oceans[7]. The authors also relate this temperature change to a burst of faunal turnover in North America in the early Oligocene, noting that the wave of extinctions occurred primarily among cool-blooded animals that might have been most sensitive to a rapid decline in temperature. In further analysing their data using other isotope–climate relationships, however, the authors find no evidence for the changes in seasonal temperature range or aridity that have been inferred in other studies[8,9].

Together, these two studies[2,3] convincingly demonstrate the global nature of climate and biotic change during the early Oligocene. They strongly support challenges to the ocean-circulation hypothesis for the early Oligocene event[10,11], and imply that a global forcing agent, such as declining concentrations of atmospheric greenhouse gases[12], played a role in triggering the build-up of Antarctic ice. But the detailed links between the observed climate changes and global forcing remain unresolved.

For example, Zanazzi and colleagues' lack of evidence for changes in aridity accompanying the major cooling of the North American continent is puzzling, both because of the range of existing evidence for aridification[8], and because continental cooling should have decreased the atmosphere's capacity to carry water to continental interiors. However, although their data offer no obvious support for Northern Hemisphere aridification, the interpretation of these data is less well constrained than for their coupled tooth/bone isotope measurements, and they may not be entirely inconsistent with the changes noted in earlier studies.

In Asia, climate-model simulations are needed to test whether climate cooling resulting from reduced CO_2 levels can indeed explain drying in this region. If they cannot, it is possible that Asian drying and early Oligocene climate change elsewhere are indirectly linked. One possibility, which remains to be tested, is that regional tectonic events (that is, uplift of the Himalayas) caused both the observed aridification in central Asia and far-afield climate change in Antarctica and North America. A pulse of Himalayan uplift near the Eocene–Oligocene boundary could have caused drying in central Asia by creating a rain shadow or increasing the seasonality of precipitation in the region. It might simultaneously have triggered a global reduction in atmospheric CO_2 concentrations through increased weathering of the mountains, which pulls CO_2 from the atmosphere.

What is clear is that these two studies[2,3] show that the early Oligocene was a turning point not just for Antarctica, but for Earth as a whole. If that event can serve as a guide, more than just polar ice is at stake as atmospheric CO_2 levels push the planet's climate towards an increasingly warm, potentially ice-free state. ■

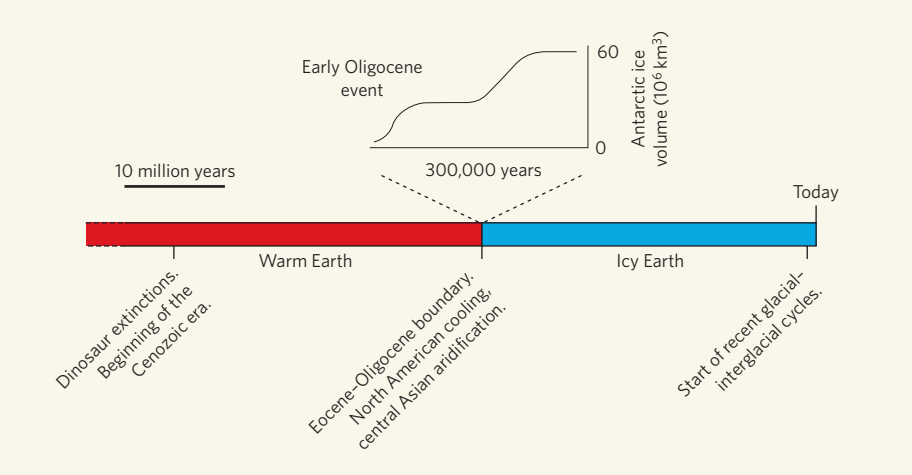


Figure 1 | The Eocene–Oligocene boundary in context. This timeline shows the transition from the warm pre-Oligocene Earth to today's icy planet that occurred at the Eocene–Oligocene boundary (33.7 million years ago), approximately midway between the extinction of the dinosaurs and today. The drying in central Asia and cooling of North America described by Dupont-Nivet *et al.*[2] and Zanazzi *et al.*[3] coincided with an abrupt, 300,000-year build-up of continental-scale ice sheets on Antarctica. (Graphic after ref. 1.)

Gabriel J. Bowen is in the Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana 47907, USA.
e-mail: gjbowen@purdue.edu

1. Coxall, H. K., Wilson, P. A., Pälike, H. & Lear, C. H. *Nature* **433,** 53–57 (2005).
2. Dupont-Nivet, G. *et al. Nature* **445,** 635–638 (2007).
3. Zanazzi, A., Kohn, M. J., MacFadden, B. J. & Terry, D. O. Jr *Nature* **445,** 639–642 (2007).
4. Kennett, J. P. *J. Geophys. Res.* **87,** 3843–3860 (1977).
5. Sijp, W. P. & England, M. H. *J. Phys. Oceanogr.* **34,** 1254–1266 (2004).
6. Meng, J. & McKenna, M. C. *Nature* **394,** 364–367 (1998).
7. Lear, C. H., Elderfield, H. & Wilson, P. A. *Science* **287,** 269–287 (2000).
8. Prothero, D. R. & Berggren, W. A. (eds) *Eocene–Oligocene Climatic and Biotic Evolution* (Princeton Univ. Press, 1992).
9. Ivany, L. C., Patterson, W. P. & Lohmann, K. C. *Nature* **407,** 887–890 (2000).
10. Huber, M. *et al. Paleoceanography* **19,** 1–12 (2004).
11. DeConto, R. M. & Pollard, D. *Nature* **421,** 245–249 (2003).
12. Pagani, M., Zachos, J. C., Freeman, K. H., Tipple, B. & Bohaty, S. *Science* **309,** 600–603 (2005).

STRUCTURAL BIOLOGY

Molecular machinery in action

Ad Bax and Dennis A. Torchia

Nuclear magnetic resonance is the best way to study motion in proteins, but it could be applied only to small systems. This limitation has been overcome to reveal the dynamics of a large protein complex.

If biology were a car, structural biologists would be looking under the bonnet to find out how the engine works. Put more prosaically, structural biology aims to understand how biology works at the molecular level. Much information is gleaned by studying at atomic resolution the three-dimensional structures of molecules that make up living organisms, and the interactions of these molecules with one another.

But, just as a photograph of a car engine doesn't tell us how that engine actually works, static biological structures alone may not suffice for us to work out the functional mechanisms of biomolecules. Instead, a video of the structural dynamics — analogous to observations of a running car engine made from inside and out — would be much more revealing. Nuclear magnetic resonance (NMR) spectroscopy provides just such an observational tool, allowing precise measurement of both the range of molecular motion and the speed at which elements within biological structures move. On page 618 of this issue, Sprangers and Kay[1] report their use of an original NMR technique to garner details of atomic motions in a molecular 'machine': the 20S proteasome core particle (CP)*. With a molecular mass of 670 kilodaltons (kDa), this protein complex is much larger and more complicated than anything previously studied at this level of detail with NMR. The study provides fresh insight into how the proteasome performs its function of removing damaged and misfolded proteins from cells.

So far, the most common use of NMR in structural biology has been the determination of three-dimensional structures in solution. For this purpose, NMR has the advantage over X-ray crystallography in that it can be applied to systems that cannot be crystallized, of which there are many. However, X-ray crystallography is applicable to large molecular assemblies — such as intact viruses or the ribosome complexes responsible for protein assembly — that NMR has been unable to tackle.

The NMR frequency of any given atomic nucleus depends on the local dipolar magnetic fields from other nearby nuclei, and the orientation of its surrounding electron cloud relative to the magnetic field that is applied during an NMR experiment. To obtain a sharp spectrum, solution NMR spectroscopy relies on the rapid brownian tumbling of molecules to average the orientation-dependent interactions to zero over time. But the larger the molecule, the slower it tumbles, and the poorer the resolution in the NMR spectrum. In practice, this limits detailed structural NMR studies to systems smaller than about 50 kDa. The motion-dependent fluctuations of the local magnetic fields cause nuclear-spin magnetization to return to its natural equilibrium state, in a process known as relaxation. For molecules with internal flexibility, relaxation rates are governed not only by the overall tumbling, but also by the magnitude and rate of motions within the macromolecule.

Almost 10 years ago it was demonstrated that, although relaxation-causing magnetic interactions are generally additive, they can also interfere with one another[2]. This interference can lead to greatly enhanced resolution for 'backbone' amide signals — that is, for the NMR signals from the amide bonds that connect the amino acids in a protein — using a procedure known as transverse relaxation optimized spectroscopy (TROSY). These amides are not only the cornerstones for determining protein structures by NMR, they also provide information about protein molecular motion.

However, it is the side chains of proteins that interact most frequently with other molecules, so how can information be obtained about these peripheral groups? During the past few years, Kay and co-workers have developed an analogous TROSY technology that focuses on methyl (CH_3) groups[3,4]. This technique is even more powerful than the original amide TROSY method, and results in sharp NMR signals for molecules exceeding 1 MDa in size. Methyl-TROSY also takes advantage of the fact that methyl groups, such as those found in certain amino-acid side chains, spin very rapidly about their axes of three-fold symmetry, so scaling down their effective magnetic interactions before overall tumbling.

Sprangers and Kay[1] now use this technique to examine the 20S proteasome CP. The three-dimensional structure of this complex, which had previously been elucidated by X-ray crystallography, is constructed from two types of protein building block, referred to as α and β. Each type of building block forms two heptameric rings (α_7 and β_7), which stack together in an α_7–β_7–β_7–α_7 manner to form the barrel-shaped CP (see Fig. 3c of ref. 1). The two outer α_7 rings form the entrance channels for substrates and the binding sites for other multimeric protein complexes, including that of the heptameric 11S activator, which modulates proteasomal activity. The interfaces between the α_7 and β_7 rings enclose two antechambers where protein substrates are stored before degradation, and the two inner β_7 rings enclose the catalytic chamber where the active sites for degradation are sequestered within the lumen of the barrel.

The authors use methyl-TROSY to study the motions of almost 100 methyl groups that are found in isoleucine, leucine and valine amino-acid residues in the α-subunits. They do this for the α-subunit in isolation, for assembled α_7 rings and for the full CP, and also examine the effect of binding of the 11S activator to the CP. In this way, they identify rapid, large-amplitude motions for a region of the α-protein implicated in transport of the proteasome to cell nuclei in mammalian systems. These movements occur on a time scale that is fast, compared with the nearly 200 nanoseconds it takes the CP to change its orientation by one radian as it tumbles. Much slower motions (of the order of a millisecond) are seen for methyl groups lining the inside of the β-barrel. These all occur at the same rate, which strongly suggests that they are correlated with one another and that they form part of a unique mechanistic process — presumably the transport of substrate to the catalytic chamber.

The remarkable success of Sprangers and Kay in obtaining high-quality NMR data for such a large system derives in part from their careful choice of a highly symmetrical complex, in which all of the subunits of the same type (either α or β) are magnetically equivalent. The CP complex can be dissected into stable components, which can then be studied in isolation and as reassembled particles. This feature was essential to assigning each of the almost 100 simultaneously detected NMR signals to its specific methyl group. However, the need for high symmetry will be reduced by the substantial increase in sensitivity afforded by the next generation of NMR machines. The prospects are therefore excellent that NMR will provide information about the internal dynamics and transient interactions of a wide variety of molecular machines. This information, when combined with structural data from X-ray diffraction and electron microscopy, will guide our understanding of how biomolecules work at a molecular level. ■

Ad Bax and Dennis A. Torchia are at the National Institutes of Health, Bethesda, Maryland 20892, USA.
e-mails: bax@nih.gov; dtorchia@mail.nih.gov

1. Sprangers, R. & Kay, L. E. *Nature* **445,** 618–622 (2007).
2. Pervushin, K., Riek, R., Wider, G. & Wuthrich, K. *Proc. Natl Acad. Sci. USA* **94,** 12366–12371 (1997).
3. Sprangers, R., Gribun, A., Hwang, P. M., Houry, W. A. & Kay, L. E. *Proc. Natl Acad. Sci. USA* **102,** 16678–16683 (2005).
4. Tugarinov, V., Hwang, P. M. & Kay, L. E. *Annu. Rev. Biochem.* **73,** 107–146 (2004).

*This article and the paper concerned[1] were published online on 21 January 2007.

NEWS & VIEWS FEATURE

EVOLUTIONARY BIOLOGY

Out of thin air

John F. Allen and William Martin

The invention of oxygenic photosynthesis was a small step for a bacterium, but a giant leap for biology and geochemistry. So when and how did cells first learn to split water to make oxygen gas?

The oxygen that gives us the breath of life is renewed by sunlight falling on plants, algae and a particular class of bacterium called cyanobacteria — all of which produce molecular oxygen (O_2) as a waste product of photosynthesis (Box 1)[1]. Biologists agree that cyanobacteria invented the art of making oxygen[2], but when and how this came about remain uncertain.

Oxygenic photosynthesis involves about 100 proteins that are highly ordered within the photosynthetic membranes of the cell. The main players are two molecular machines, photosystem I and photosystem II, that act as electrochemical solar cells. With the help of chlorophyll (the pigment that makes plants green), they transform sunlight into electrical current (Fig. 1). Photosystem II generates an electrochemical potential of + 1.1 volts, enough to remove two electrons from each of two water molecules, making a molecule of O_2 at a cost of four photons — one for each electron moved. Photosystem II performs this remarkable feat only when photosystem I is present to dispose of the electrons. Photosystem I grabs the four electrons and uses four more photons to deposit them, in two pairs, on an electron carrier called $NADP^+$. $NADP^+$ ultimately transfers the electrons to carbon dioxide, thereby providing the energy to make carbon-based sugars and the other molecules of life: light makes life and oxygen out of water and thin air.

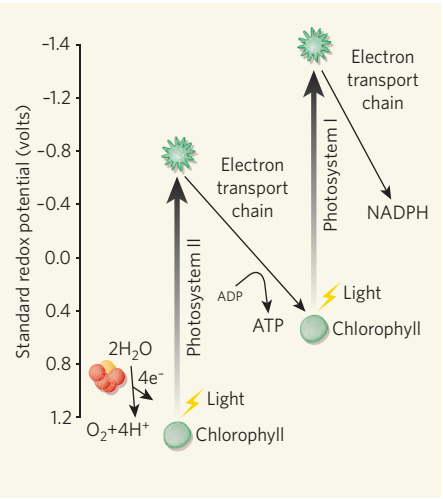


Figure 1 | The two photosystems in photosynthesis. Photosystems I and II absorb light energy, convert it into electrochemical potential, and are connected in series electrically. These two 'light reactions' of photosynthesis form links in a chain of electron (e^-) transfers that is coupled, by means of proton pumping, to synthesis of the energy-storage molecule adenosine triphosphate (ATP). The electron transport chain of photosynthesis, also known as the Hill and Bendall Z-scheme, ends with photosystem I delivering electrons to $NADP^+$, making NADPH. ATP and NADPH drive the 'dark reactions' that transfer the electrons to CO_2 so as to provide the energy to make sugars and the other molecules of life. The chain begins when water is oxidized to oxygen by the very high electrochemical potential of photosystem II. The catalyst of water oxidation (Box 3) is shown here as a cluster of four red spheres and one yellow one.

A matter of time

Oxygenic photosynthesis is the only significant source of O_2 known, so indications for O_2 in the geological record should date its origin. Geochemical evidence suggests that 2.3 billion years ago atmospheric oxygen had risen to more than 10^{-5} of its present concentration[3]. So this date should set a minimum age for cyanobacteria[4]. But might they be older? Combining palaeontology and molecular phylogeny, Tomitani *et al.*[5] claim that they are, arguing that 2.3 billion years maps not to the infancy, but to the middle ages of cyanobacterial evolution. Their argument is based on the presence in certain cyanobacteria of modified cells (heterocysts) that are dedicated to nitrogen fixation. Heterocysts occur only in the more specialized cyanobacterial lineages and offer a virtually O_2-free environment for nitrogen fixation, which is inhibited by O_2. Tomitani *et al.* reason that heterocysts would not have been needed until the advent of atmospheric O_2, and that their origin coincided with the onset of atmospheric O_2 accumulation.

Could cyanobacteria date back to 3.8 billion years ago, or even earlier? Not according to the geochemical data from a 3.4-billion-year-old sedimentary formation in South Africa called the Buck Reef Chert[6]. This massive formation shows evidence of having been deposited through carbon fixation by photosynthesizing organisms, but it contains none of the mineral traces expected from the production of O_2. Thus, the simplest interpretation is that the deposit was laid down by anoxygenic photosynthesis for which the source of electrons was molecular hydrogen or other inorganic

Box 1 | The discovery of photosynthesis

In 1772, Joseph Priestley described how oxygen is consumed by combustion or by respiration using a burning candle, or a live mouse, in a closed glass jar. His discovery predated the term oxygen, which was coined by Lavoisier[20,21]. The results were reported in a paper entitled 'Observations on Different Kinds of Air'[1]:

"...I flatter myself that I have accidentally hit upon a method of restoring air which has been injured by the burning of candles, and that I have discovered at least one of the restoratives which nature employs for this purpose. It is vegetation. In what manner this process in nature operates, to produce so remarkable an effect, I do not pretend to have discovered; but a number of facts declare in favour of this hypothesis...

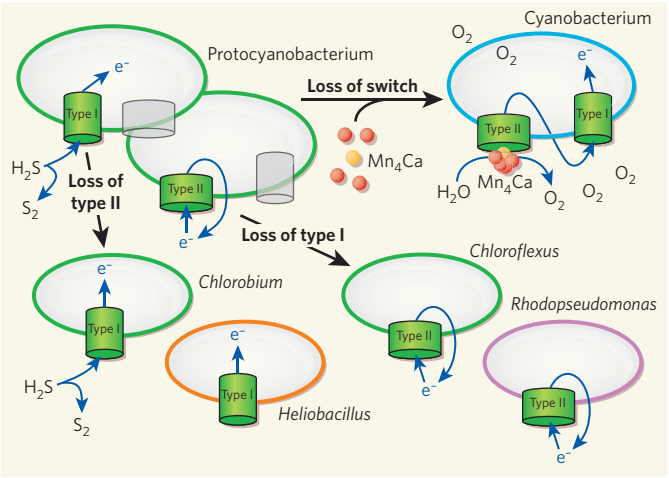
"One might have imagined that, since common air is necessary to vegetable, as well as to animal life, both plants and animal had affected it in the same manner, and I own that I had that expectation, when I first put a sprig of mint into a glass-jar, standing inverted in a vessel of water; but when it had continued growing there for some months, I found that the air would neither extinguish a candle, nor was it at all inconvenient to a mouse, which I put into it.

"...Accordingly, on the 17th of August 1771, I put a sprig of mint into a quantity of air, in which a wax candle had burned out, and found that, on the 27th of the same month, another candle burned perfectly well in it. This experiment I repeated, without least variation in the event, not less than eight or ten times in the remainder of the summer." **J.F.A. & W.M.**

Box 2 | A missing link in the evolution of photosynthesis

A modern anaerobic 'protocyanobacterium' possessing the genes encoding both photosystems involved in photosynthesis, but expressing them differentially, would be a bona fide missing link in the evolution of this vital process. This missing link (top left) would have the ability to switch between a *Chlorobium*-like, type-I photosystem and a *Rhodopseudomonas*-like, type-II photosystem[22]. In the protocyanobacterium, if either photosystem is permanently switched off, its mutation incurs no selective penalty, and its genes are eventually lost. Such a process may have given rise to the familiar, anoxygenic species, represented in the diagram by *Chlorobium* and *Heliobacillus* (type-I photosystems), and *Chloroflexus* and *Rhodopseudomonas* (type-II photosystems).

When both photosystems are being selected for, but under different growth conditions, a mutation in the switch could cause the two photosystems to coexist in the same membrane at the same time. This is usually harmful, and selected against, because the type-I and type-II electron (e^-) transport pathways (blue arrows) share components of the rest of the electron transport chain while being linear and cyclic, respectively, so the two photosystems would interfere with each other.



However, with a catalyst that oxidizes water (producing oxygen and electrons), the situation would be transformed because a continuous, smooth, linear electron flow is assured; the two photosystems would become complementary, and together begin to use light to drive oxygen production by the first true cyanobacterium (top right). The catalyst of water oxidation (Box 3) is shown here as four red spheres (each representing a manganese (Mn) atom) and one yellow one (representing a calcium (Ca) atom).

So, is there anything like a protocyanobacterium around today? One true cyanobacterium, *Oscillatoria limnetica*, turns off its genes for photosystem II in the presence of H_2S, and thus 'reverts' from oxygenic photosynthesis[23] to the kind of anoxygenic photosynthesis seen in *Chlorobium*. Other cyanobacteria retain the ability to use thiosulphate as an electron donor, in place of water, and photosystem II[24]. But the converse switch, turning off photosystem I so as to survive like *Rhodopseudomonas*, has not been seen. *Chloroflexus aurantiacus* has a type-II photosystem in combination with a peculiar kind of light-collecting antenna, called a chlorosome, that is otherwise specific to the type-I-containing *Chlorobium*. This has led to the suggestion that *Chloroflexus* is descended from a protocyanobacterium that lost its type-I photosystem[25]. **J.F.A. & W.M.**

molecules[4,6]. If 3.4 billion years is taken as an upper boundary for the age of O_2 synthesis, the origin of oxygenic photosynthesis would fall within the range of 3.4 billion to 2.3 billion years ago. This range is admittedly rather imprecise, but it is something. Of course, absence of evidence is not evidence of absence, and other authors suggest that O_2 could have been produced as early as 4.0 billion years ago, but was rapidly consumed[7,8].

On the double

So how did oxygen production originate? Some modern bacteria carry out anoxygenic photosynthesis, which is assumed to have been the precursor of the oxygen-producing reaction. But these bacteria use either a protein complex similar to photosystem I, or one similar to photosystem II — never both systems together[9]. For example, *Rhodopseudomonas palustris* is one of the purple photosynthetic bacteria, which possess only a photosystem-II-like reaction centre that drives a cyclic electron transport pathway (Box 2). Purple bacteria are versatile organisms, able to grow in the dark using cellular respiration as a source of energy, or in the light by switching on their photosystem. *Chlorobium tepidum,* on the other hand, is an anaerobic green bacterium that uses only a photosystem-I-like reaction centre to harness light energy, drawing electrons from hydrogen sulphide for linear electron transport.

No tree of bacterial life can readily account for the observed distributions of the two sets of photosystem genes among the species[10]. This has left biologists with little alternative but to suggest that genes encoding the photosystems have moved across species boundaries during evolution, a process called lateral gene transfer. But transfer from what to what? Did a cell containing photosystem I subsequently acquire photosystem II (ref. 11) or vice versa[12]? Or did both photosystems arise in a precursor of cyanobacteria — a 'protocyanobacterium' — later to be exported to other lineages by lateral transfer (Box 2)?

Mulkidjanian and colleagues[13] recently weighed in with evidence for the latter alternative. These authors compared the genome sequences of several cyanobacteria and anoxygenic photosynthetic bacteria, and identified a core set of genes involved in photosynthesis. They suggest that these genes arose in a now-extinct group of anoxygenic bacteria and then moved between species to other distant lineages. Like others[11], they suggest[13] that photosystem I was the ancestral prototype, from which an evolutionary precursor of photosystem II arose. The reaction-centre cores of the photosystems are similar in structure[14], and their divergence probably began with a simple duplication of the associated gene cluster. But before the water-splitting complex evolved, what would a bacterium with two different and specialized, non-oxygenic photosystems have done with them? Probably just what modern bacteria do: express them when needed, with the help of a regulatory switch.

How would a putative photosystem-switching bacterium then smooth the evolutionary path to oxygenic photosynthesis? It would have been only a small step away from the cyanobacterial state of oxygenic photosynthesis (Box 2), provided that it underwent the right mutation — disabling the regulatory switch — and provided that this happened in the right environmental setting and at the right time.

Centre stage

Of course, any proposed evolutionary scheme must explain not only the presence of two photosystems, but also how the oxidation of water came to be coupled to their reactions. This essential catalysis in photosynthetic oxygen production is carried out by metal ions: four oxidized manganese (Mn) atoms and a calcium (Ca) atom. These atoms are held in place by proteins, at a site called the water-splitting complex of photosystem II. One could say that the job of photosystem II in oxygenic photosynthesis is not to oxidize water, but rather to extract, one at a time, four electrons from the Mn_4Ca cluster. Only when this cluster is fully four electrons short does it replace them. It grabs them back in one fell swoop from two water molecules, thereby restoring its original charge state and releasing, in the process, the most useful waste product ever known — O_2.

The Mn_4Ca cluster lies on the surface of photosystem II that faces the environment. Its precise structure and the chemistry of its catalytic reactions have long been sought, not least because of the prospect of mimicking nature to exploit sunlight as a clean and renewable source of energy. At the end of last year, Yano

et al.[15] brought the cluster sharply into focus (Box 3). Their high-resolution structure shows, as some expected[8,16], a similarity to manganese oxide minerals. Might the Mn_4Ca cluster carry an evolutionary imprint of the specific environment where water-splitting arose? And is there anything special about manganese that makes it nature's only solution to water splitting?

It has been known for some time that manganese ions of the type that occur in the water-splitting complex are readily oxidized by ultraviolet light of wavelength less than 240 nanometres[8,17]. In this photooxidation process, ultraviolet light is absorbed by the electron cloud of a manganese atom, causing it to throw off an electron, and leaving the Mn atom in a more oxidized state. A similar photooxidation, but one involving chlorophyll and visible light, occurs in photosynthetic reaction centres. Moreover, it is well known that ultraviolet light damages the photosystem II of modern cyanobacteria through a process called photoinhibition. This involves absorption of ultraviolet light by Mn atoms in the water-splitting complex, which causes them to break away from photosystem II, depriving it of the ability to oxidize water[18]. So the light-absorbing behaviour of Mn ions is a biologically relevant process in modern photosynthesis.

Two into one

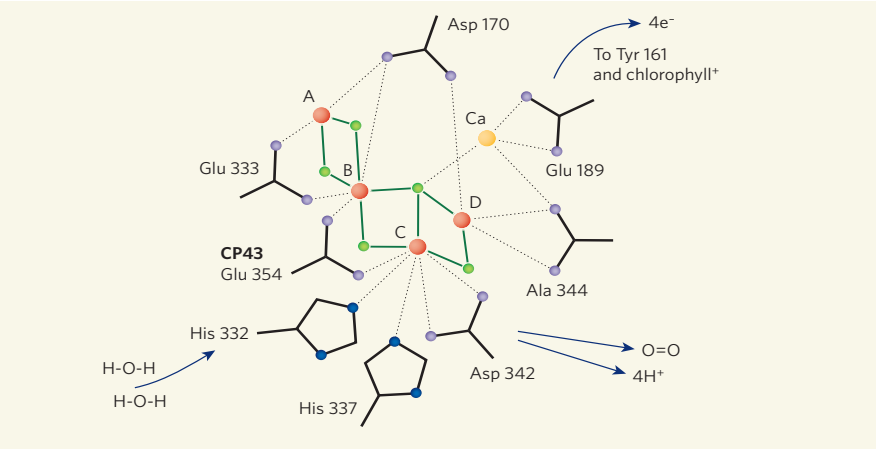
Given that there was no ozone layer to filter out the Sun's ultraviolet light before the origin of water splitting, what might have happened if an organism possessing both photosystems, but expressing only photosystem II, was introduced into an ancient, aquatic, Mn-containing environment? Photooxidation of environmental Mn would 'push' electrons into photooxidized chlorophyll of photosystem II. That would have been deadly, as a log-jam of electrons would stall the smoothly running electron transport cycle — unless the protocyanobacterium responded, either physiologically or through mutation, by expressing photosystem I. This complex would bleed off the surplus electrons from photosystem II, and create precisely the flow of electrons seen in cyanobacteria today: a linear flow through two photosystems, and from manganese to carbon dioxide (Box 2). The circuit would hardly have been perfect to start with, but the basic wiring would have been right. Notably, the Mn atoms of the water-splitting complex are bound directly to the proteins of the photosystem II reaction centre, without an intervening protein or electron carrier. This suggests that no major evolutionary invention was required for photosystem II to tap environmental Mn as an electron source.

The final step to the evolution of oxygen would then have entailed the transition from a photosystem exploiting an environmental supply of soluble manganese, where each electron-donating Mn ion reached photosystem II by simple diffusion, to one in which four manganese atoms (and a calcium of as-yet-unknown function) were held in place. A fine-tuning of photosystem II by natural selection to optimize its reduction/oxidation potential[19] would allow it to oxidize its biologically portable manganese reservoir four times in a row.

The best evidence for this evolutionary scheme would be the discovery of a modern-day protocyanobacterium. Although it is possible that all protocyanobacterial lineages have died out, we prefer to think that the missing link is still out there. An organism that possesses and expresses two photosystems, but only one at a time, may have adapted to present-day environments, such as eutrophic, anaerobic lakes with intermittent sulphide influx. An anoxygenic phototroph that switches gene expression between type-I and type-II photosynthesis could still be with us, unchanged from the time when its ancestors gave birth to our familiar, aerobic world[20]. ■

John F. Allen is at the School of Biological and Chemical Sciences, Queen Mary, University of London, Mile End Road, London E1 4NS, UK.
e-mail: j.f.allen@qmul.ac.uk
William Martin is at the Institut für Botanik III, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany.
e-mail: w.martin@uni-duesseldorf.de

Box 3 | The water-splitting reaction centre



A model of the structure of the Mn_4Ca cluster in photosystem II (structure reproduced with permission from ref. 15). The 'model II' structure, from X-ray spectroscopy of a crystal of the photosystem II reaction centre, shows four Mn atoms (red; A–D) and a Ca atom (yellow; Ca). The figure shows a prediction of how the water-splitting metal-atom cluster fits into the highest-resolution protein structure so far obtained with X-ray crystallography. Predicted non-covalent bonds are shown as dashed lines. The bonds between Mn and O atoms are green. Protein amino-acid side-chains are identified by their three-letter abbreviation and number in the sequence of the 'D_1' reaction-centre protein. One of the side-chains is provided instead by glutamate 354 of a protein called 'CP43'. The Mn_4Ca cluster occupies a position close to the outer surface of the photosynthetic membrane, which would be below the plane of the graphic. When four electrons have been removed, one from each manganese atom, two water molecules become oxidized to give one diatomic oxygen molecule. Water enters and oxygen leaves at the outer surface of the membrane. Each electron is transferred in turn from the cluster to a tyrosine side-chain (Tyr 161; not shown) and moves through the reaction centre of photosystem II to a very strong electron acceptor deep in the reaction centre — the chlorophyll molecule that has been oxidized by light. **J.F.A. & W.M.**

1. Priestley, J. *Phil. Trans.* **62,** 147–264 (1772).
2. Blankenship, R. E. *Molecular Mechanisms of Photosynthesis* (Blackwell Science, Oxford, 2002).
3. Bekker, A. *et al. Nature* **427,** 117–120 (2004).
4. Olson, J. M. *Photosynth. Res.* **88,** 109–117 (2006).
5. Tomitani, A., Knoll, A. H., Cavanaugh, C. M. & Ohno, T. *Proc. Natl Acad. Sci. USA* **103,** 5442–5447 (2006).
6. Tice, M. M. & Lowe, D. R. *Geology* **34,** 37–40 (2006).
7. Kasting, J. F. *Nature* **443,** 643–645 (2006).
8. Russell, M. J. & Hall, A. J. in *Evolution of Early Earth's Atmosphere, Hydrosphere, and Biosphere — Constraints from Ore Deposits* (eds Kesler, S. E. & Ohmoto, H.) 1–32 (Geol. Soc. Am., Boulder, CO, 2006).
9. Bryant, D. A. & Frigaard, N.-U. *Trends Microbiol.* **14,** 488–496 (2006).
10. Raymond, J., Zhaxybayeva, O., Gogarten, J. P. & Blankenship, R. E. *Phil. Trans. R. Soc. Lond. B* **358,** 223–230 (2003).
11. Baymann, F., Brugna, M., Muhlenhoff, U. & Nitschke, W. *Biochim. Biophys. Acta* **1507,** 291–310 (2001).
12. Xiong, J. *et al. Science* **289,** 1724–1730 (2000).
13. Mulkidjanian, A. Y. *et al. Proc. Natl Acad. Sci. USA* **103,** 13126–13131 (2006).
14. Schubert, W. D. *et al. J. Mol. Biol.* **280,** 297–314 (1998).
15. Yano, J. *et al. Science* **314,** 821–825 (2006).
16. Sauer, K. & Yachandra, V. K. *Proc. Natl Acad. Sci. USA* **99,** 8631–8636 (2002).
17. Anbar, A. D. & Holland, H. D. *Geochim. Cosmochim. Acta* **56,** 2595–2603 (1992).
18. Hakala, M. *et al. J. Exp. Bot.* **57,** 1809–1816 (2006).
19. Rutherford, A. W. & Faller, P. *Phil. Trans. R. Soc. Lond. B* **358,** 245–253 (2003).
20. Lane, N. *Oxygen: The Molecule that Made the World* (Oxford Univ. Press, 2002).
21. Govindjee & Krogmann, D. in *Discoveries in Photosynthesis* (eds Govindjee, Beatty, J. T., Gest, H. & Allen, J. F.) 63–105 (Springer, Dordrecht, 2005).
22. Allen, J. F. *FEBS Lett.* **579,** 963–968 (2005).
23. Oren, A. & Padan, E. *J. Bacteriol.* **133,** 558–563 (1978).
24. Koenig, F. *Bot. Acta* **103,** 54–61 (1990).
25. Allen, J. F. & Puthiyaveetil, S. in *Photosynthesis: Fundamental Apects to Global Perspectives* (eds van der Est, A. & Bruce, D.) 753–756 (Alliance Communications Group, Lawrence, KS, 2005).

VETERINARY EPIDEMIOLOGY

Vaccination strategies for foot-and-mouth disease

Arising from: M. J. Tildesley *et al. Nature* **440**, 83–86 (2006)

When foot-and-mouth disease struck the United Kingdom in 2001, the traditional 'stamping out' policy of 1967–68 was supplemented by the pre-emptive culling of animals in premises contiguous to infected premises. A model proposed by Tildesley *et al.*[1] indicates that the introduction of vaccination should at least halve the number of premises that would need to be subjected to culling in the event of another outbreak. We contest, however, that the overlapping confidence intervals of the outputs of their model, and the inconsistency of their results compared with those from previous models, call into question the model's value as a decision tool, while adding little to the recognized tenet of ring vaccination.

Foot-and-mouth disease occurred regularly in the United Kingdom over the first two-thirds of the twentieth century[2]. As a result, a 'stamping out' policy was developed that was based on movement restrictions, enforcement of biosecurity, rapid slaughter of animals from infected premises, and tracing and slaughter of dangerous contacts. A disease-free period of more than 30 years followed the epidemic of 1967–68. When the disease returned in 2001, this 'stamping out' policy was augmented by pre-emptive culling in premises contiguous to infected premises; vaccination was not used.

Tildesley *et al.*[1] model 'reactive' vaccination as an alternative to pre-emptive culling on contiguous premises and find that this strategy would at least halve the number of premises for culling. They explore different strategies and conclude that the impact of an epidemic can be restricted most effectively by using reactive ring vaccination, prioritizing farms closest to infected premises. The authors claim that this technique would be more effective than the prioritization of high-risk farms. They also conclude that the addition of culling in contiguous premises is not advantageous when vaccination is practised. However, the overlapping confidence intervals of the parameters associated with the impact of the epidemic under different strategies (Table 2 of ref. 1) indicate that the model cannot offer a clear choice between different scenarios, making its value as a decision tool questionable.

The value of this modelling approach in decision support is further undermined because the conclusions of previous papers based on the same model[3,4] are inconsistent with those of Tildesley *et al.*[1]. For example, Keeling *et al.*[3] showed that ring vaccination of cattle would reduce the total cull by only about 15%, and that removing culling in contiguous premises from a control strategy that includes ring vaccination would have doubled the total cull; Keeling *et al.*[4] also pointed out that, although ring vaccination generates some benefit, even 10-km rings reduce the size of the epidemic by only about 20%, and that "vaccinating the nearest farms [to infected premises] would be futile", suggesting that reactive vaccination should be targeted to farms that are predicted to contribute most to future spatial transmission of disease.

The model used by the authors, here[1] and previously[3,4], is parameterized in such a way as to favour disease-control methods (culling, vaccination or both) that are pre-emptive and localized[5]. Farms are assumed to become instantly and maximally infectious four days before the simulated infection is reported. The spatial simulation of disease spread in the models is based on spread-tracing information, which the authors themselves previously recognized[3] as likely to be biased towards overestimating the amount of short-distance spread of disease. The model still needs to be validated against data not used in its construction[6]. Without this validation, it remains an untested hypothesis that has produced inconsistent and imprecise conclusions about vaccination policy, and therefore adds little, if anything, to the already well established principle of ring vaccination[7].

Richard P. Kitching*, Nicholas M. Taylor†, Michael V. Thrusfield‡
*National Centre for Foreign Animal Diseases, Winnipeg R3E 3M4, Canada
†Veterinary Epidemiology and Economics Research Unit, University of Reading, Earley Gate, Reading RG6 6AR, UK
‡Veterinary Clinical Sciences, University of Edinburgh, Easter Bush Veterinary Centre, Roslin EH25 9RG, UK
m.thrusfield@ed.ac.uk

1. Tildesley, M. J. *et al. Nature* **440,** 83–86 (2006).
2. HMSO *Report of the Committee of Enquiry on Foot-and-Mouth Disease 1968* Vols 1, 2 (Her Majesty's Stationery Office, London, 1969).
3. Keeling, M. J. *et al. Science* **294,** 813–817 (2001).
4. Keeling, M. J., Woolhouse, M. E. J., May, R. M., Davies, G. & Grenfell, B. T. *Nature* **421,** 136–142 (2003).
5. Kitching, R. P., Thrusfield, M. V. & Taylor, N. M. *Rev. Sci. Tech.* **25,** 293–311 (2006).
6. Spedding, C. R. W. *An Introduction to Agricultural Systems* 2nd edn (Elsevier Applied Science, London, 1988).
7. Donaldson, A. I. & Doel, T. R. *Vet. Rec.* **131,** 114–120 (1992).

Competing financial interests: declared none.
doi: 10.1038/nature05604

VETERINARY EPIDEMIOLOGY

Tildesley et al. *reply*

Replying to: R. P. Kitching, N. M. Taylor & M. V. Thrusfield *Nature* **445,** doi: 10.1038/nature05604 (2007)

Kitching *et al.*[1] express concerns about our mathematical models of reactive vaccination strategies for the control of foot-and-mouth disease epidemics[2]. However, in our view, these concerns are misplaced.

Kitching *et al.* note that simulated epidemic sizes can be highly variable, but they misconstrue the reason for this. The variability does not imply a lack of precision in the outputs of the model; rather, the model is capturing a well recognized feature of outbreaks of infectious diseases in general[3] and of foot-and-mouth disease in particular[4]. By performing tens of thousands of simulations, we can robustly establish the relative effectiveness of different control options on average across the whole range of expected outcomes. We regard this as a major strength of our approach: it would be most unwise to recommend control policies that were potentially vulnerable to expected stochastic variability. Our approach also allows us to consider alternative criteria for an effective control policy, such as minimizing the probability of a very large epidemic.

These results are not inconsistent with our earlier work[5,6] (Kitching *et al.* are quoting this out of context), in which we considered different scenarios — specifically, different start dates for vaccination. Previous results can be recovered readily by using the latest version of the model. The novelty of our recent study[2] lies in the finding that, in a resource-limited situation, considerable and unanticipated gains in

efficiency can be achieved by optimizing the order in which farms are vaccinated. The most efficient strategy in our model is not equivalent to conventional ring vaccination strategies. We suggest that this quite general result will turn out to be applicable to other reactive vaccination problems.

Our model does favour the implementation of local control measures, but this is an emergent property of fitting the model to data from 2001, not an in-built assumption. Kitching *et al.*[7] mention their earlier work (their ref. 5), which cites papers that question the effectiveness of localized control[8,9]. In our view, the analyses reported in those two studies are open to criticism — for example, we find their regression models inappropriate for time-series data, and their criteria for statistical significance questionable; moreover, when we repeat the analyses described, we obtain results supporting localized control.

Regarding the other points raised, we stress that the model does not make any assumptions that are inconsistent with the 2001 data[10–12]. We undertook extensive sensitivity analysis to confirm that our conclusions were robust to uncertainty in parameter values, including uncertainties about the infectiousness of farms, the efficacy of vaccines and the transmission kernel[2]. Although we did use the full 2001 data set to make best use of the available information, the underlying model had been validated previously[5] by successful prediction of phases of the 2001 epidemic. The model is, of course, a simulation tool, and not a hypothesis.

Quantitative models have an important contribution to make to the control of future epidemics of foot-and-mouth disease and other infectious diseases[13]. Although we think that our model[2] will be of value as a decision tool, it should only ever be used — as with any model — as one of various sources of information available to policy makers.

Michael J. Tildesley*, Nicholas J. Savill†, Darren J. Shaw‡, Rob Deardon§, Stephen P. Brooks||, Mark E. J. Woolhouse†, Bryan T. Grenfell¶, Matt J. Keeling*

*Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, UK
email: M.J.Keeling@warwick.ac.uk
†Centre for Infectious Diseases, University of Edinburgh, Ashworth Laboratories, Kings Buildings, Edinburgh EH9 3JT, UK
‡Veterinary Clinical Sciences, Easter Bush Veterinary Centre, Roslin, Midlothian EH25 9RG, UK
§Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario N1G 2W1, Canada
||Statistical Laboratory, Centre for Mathematical Sciences, Cambridge CB3 0WB, UK
¶Center for Infectious Disease Dynamics, Pennsylvania State University, University Park, Pennsylvania 16802, USA

1. Kitching, R. P., Taylor, N. M. & Thusfield, M. V. *Nature* **445,** doi: 10.1038/nature05604 (2007).
2. Tildesley, M. J. *et al. Nature* **440,** 83–86 (2006).
3. Becker, N. G. *Analysis of Infectious Disease Data* (Chapman & Hall, 1989).
4. Matthews, L. & Woolhouse, M. *Nature Rev. Microbiol.* **3,** 529–536 (2005).
5. Keeling, M. J. *et al. Science* **294,** 813–817 (2001).
6. Keeling, M. J., Woolhouse, M. E. J., May, R. M., Davies, G. & Grenfell, B. T. *Nature* **421,** 136–142 (2003).
7. Kitching, R. P., Thrusfield, M. V. & Taylor, N. M. *Rev. Sci. Tech.* **25,** 293–311 (2006).
8. Honhold, N., Taylor, N. M., Mansley, L. M. & Paterson, A. D. *Vet. Rec.* **155,** 287–294 (2004).
9. Thrusfield, M., Mansley, L., Dunlop, P., Pawson, A. & Taylor, J. *Vet. Rec.* **156,** 269–278 (2005).
10. Savill, N. J. *et al. J. R. Soc. Interface* (in the press); doi: 10.1098/rsif.2006.0178).
11. Hughes, G. J. *et al. J. Gen. Virol.* **83,** 1907–1914 (2002).
12. Savill, N. J. *et al. BMC Vet. Res.* **2,** 3 (2006).
13. The Royal Society *Infectious Diseases in Livestock* (The Royal Society, London, 2002).

doi: 10.1038/nature05605

ARTICLES

Axial patterning in cephalochordates and the evolution of the organizer

Jr-Kai Yu[1,4], Yutaka Satou[2], Nicholas D. Holland[1], Tadasu Shin-I[3], Yuji Kohara[3], Noriyuki Satoh[2], Marianne Bronner-Fraser[4] & Linda Z. Holland[1]

The organizer of the vertebrate gastrula is an important signalling centre that induces and patterns dorsal axial structures. Although a topic of long-standing interest, the evolutionary origin of the organizer remains unclear. Here we show that the gastrula of the cephalochordate amphioxus expresses dorsal/ventral (D/V) patterning genes (for example, bone morphogenetic proteins (*BMP*s), *Nodal* and their antagonists) in patterns reminiscent of those of their vertebrate orthlogues, and that amphioxus embryos, like those of vertebrates, are ventralized by exogenous BMP protein. In addition, Wnt-antagonists (for example, *Dkk*s and *sFRP2-like*) are expressed anteriorly, whereas *Wnt* genes themselves are expressed posteriorly, consistent with a role for Wnt signalling in anterior/posterior (A/P) patterning. These results suggest evolutionary conservation of the mechanisms for both D/V and A/P patterning of the early gastrula. In light of recent phylogenetic analyses placing cephalochordates basally in the chordate lineage, we propose that separate signalling centres for patterning the D/V and A/P axes may be an ancestral chordate character.

The evolutionary origin of the vertebrate organizer (that is, 'Spemann's organizer' in amphibians or the 'node' in birds and mammals) has been the subject of considerable speculation[1,2]. On the one hand, it has been proposed that the organizer was not present at the phylogenetic base of the chordates[2], but on the other, it has been suggested that an organizer is present in all chordates[1]. Cephalochordates (amphioxus) represent basal chordates[3–5], and have an embryonic body plan highly reminiscent of that of vertebrates. Thus, amphioxus occupies an important evolutionary position for understanding the origin of vertebrates in general and the evolutionary origin of the organizer in particular. Amphioxus eggs are small (~140 μm in diameter), have little yolk and generate a single-cell-layered blastula that undergoes simple invagination during gastrulation with little involution[6] (Fig. 1). Grafts of the dorsal lip of the amphioxus blastopore to a host blastocoel were reported to induce partial secondary axes[7], indicating that the dorsal blastopore lip of the amphioxus gastrula may be comparable to the organizer of amphibians. However, the molecular mechanisms underlying this induction are not understood.

Because the vertebrate organizer expresses a characteristic suite of genes[8–10], we asked whether the amphioxus and vertebrate organizers share homologous molecular signatures. To identify amphioxus orthologues of vertebrate organizer genes, we performed a large-scale complementary DNA analysis of the Florida amphioxus, *Branchiostoma floridae*. Five cDNA libraries (unfertilized oocytes, gastrulae, neurulae, 36 h post-fertilization larvae, and adults) were constructed and expressed sequence tags (ESTs; deposited in GenBank) were generated from 140,000 clones. BLAST searches revealed 21 orthologues of vertebrate organizer genes including several transcription factors as well as members of the BMP, Wnt and Nodal signalling pathways (Supplementary Table 1). To determine if these genes are similarly deployed in amphioxus and vertebrates, we analysed their expression during gastrulation and investigated the role of BMP protein in D/V patterning. Our results, together with

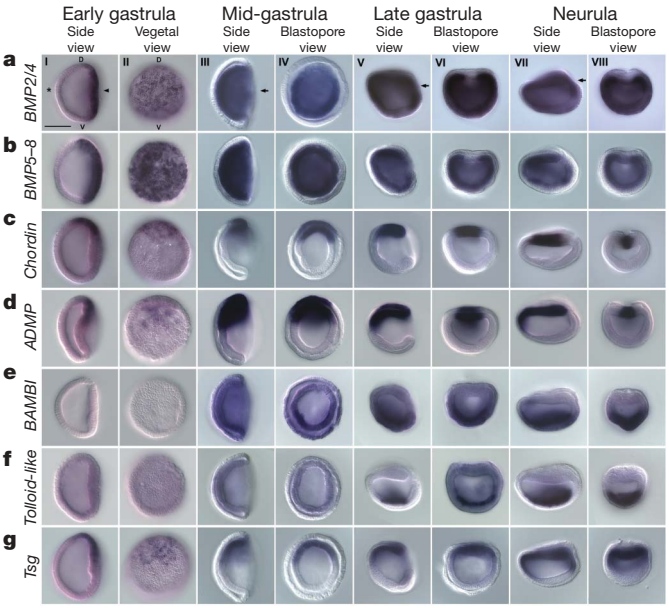


Figure 1 | Expression of genes for BMP signalling molecules during early amphioxus embryogenesis. Anterior shown at left in the side views (I, III, V, VII) and perpendicular to the plane of the page in the vegetal and blastopore views (II, IV, VI, VIII). Asterisk indicates the animal pole, which is close to the anterior pole. The arrowhead indicates the vegetal plate. By the mid-gastrula, the vegetal plate is fully invaginated. The blastopore is posterior (arrows). Scale bar, 50 μm. **a**, **b**, *BMP2/4* and *BMP5-8* are expressed throughout the mesendoderm (except the dorsal axial mesoderm). **c**, **d**, *Chordin* and *ADMP* are expressed dorsally. **e**, **f**, *BAMBI* and *Tolloid-like* are expressed ventrally. *BAMBI* expression undetectable in the early gastrula. **g**, *Tsg* is expressed dorsally. D, dorsal; V, ventral.

[1]Marine Biology Research Division, Scripps Institution of Oceanography, La Jolla, California 92037-0202, USA. [2]Department of Zoology, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan. [3]National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan. [4]Division of Biology 139-74, California Institute of Technology, Pasadena, California 91125, USA.

previous data, show not only that developmental expression of most amphioxus orthologues of vertebrate organizer genes is similar to that of their vertebrate counterparts, but also that exogenous BMP protein ventralizes amphioxus embryos as it does those of vertebrates.

Expression of BMP and Nodal signalling molecules

We isolated seven genes involved in BMP signalling (Fig. 1), which mediates D/V patterning in vertebrates; of these, all but *Twisted gastrulation* (*Tsg*) are expressed in similar patterns to their *Xenopus* orthologues (discussed below). At the onset of gastrulation, amphioxus *BMP2/4* and *BMP5-8* are co-expressed throughout the vegetal plate (that is, prospective mesendoderm). By the late gastrula, *BMP2/4* is somewhat downregulated in the ventral posterior mesendoderm and upregulated in the ectoderm, whereas both *BMPs* become downregulated in the dorsal axial mesoderm (Fig. 1a, b). The BMP antagonist *Chordin* switches on in the dorsal lip of the blastopore at the onset of gastrulation in both ectoderm and mesoderm; later its expression becomes restricted to the dorsal axial mesoderm (notochord progenitor) and the centre of the neural plate (floorplate) (Fig. 1c). Also co-expressed with *Chordin* is the gene encoding another BMP family member, anti-dorsalizing morphogenetic protein (*ADMP*), the orthologue of which is expressed in the vertebrate organizer[8,11] (Fig. 1d). In addition, we isolated an amphioxus orthologue of the BMP antagonist *Noggin* from the EST libraries (Supplementary Table 1). However, expression at the gastrula and neurula stages was undetectable by *in situ* hybridization even though riboprobes to several regions of the cDNA were used (data not shown).

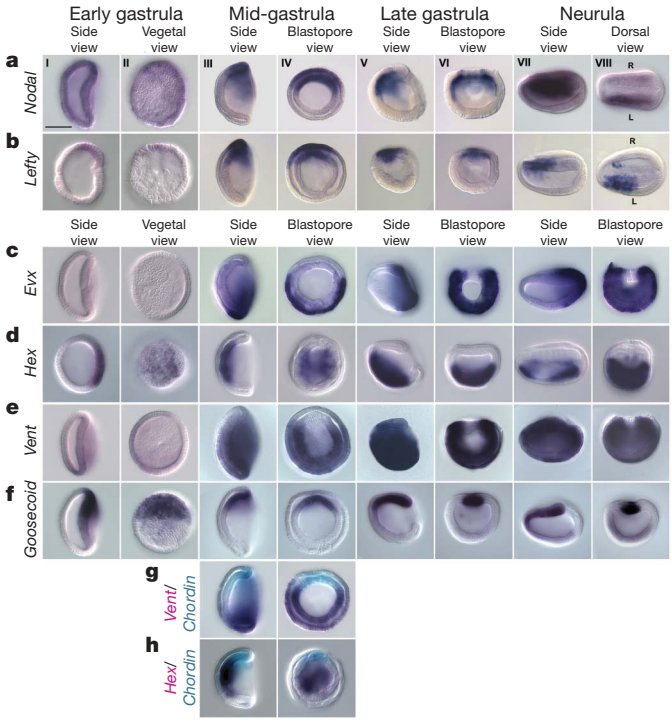


Figure 2 | Expression of *Nodal*, the *Nodal* antagonist *Lefty* and organizer-related transcription factors in early amphioxus embryos. Anterior shown towards the left in dorsal and side views. **a**, **b**, *Nodal* and *Lefty* are expressed dorsally during gastrulation. Dorsal views (column VIII) show left-sided expression of *Nodal* and *Lefty*. L, left side; R, right side. **c**, *Evx* is expressed ventrally. *Evx* expression is undetectable in early-gastrulae. **d**, *Hex* is expressed in the centre of the invaginating mesendoderm during gastrulation. **e**, *Vent* is expressed ventrally. **f**, *Goosecoid* is expressed dorsally during gastrulation. **g**, Double *in situ* hybridization showing expression of *Chordin* (cyan) and *Vent* (purple) in complementary domains. **h**, Double *in situ* hybridization showing non-overlapping expression of *Hex* (purple) and *Chordin* (cyan). Scale bar, 50 μm.

In vertebrates, the important BMP modulators *BMP and activin membrane-bound inhibitor* (*BAMBI*), and *Tsg* and *Tolloid* are co-expressed with BMPs ventrally, where they are thought to modulate BMP signalling[8]. Similarly, in amphioxus, *BAMBI* and *Tolloid-like* are expressed in the ventral two-thirds of the mid-gastrula, although *Tolloid-like* expression is limited to the ventral mesendoderm whereas *BAMBI* is also expressed in ectoderm (Fig. 1e, f). In contrast, *Tsg* is expressed in dorsal mesendoderm at the early gastrula stage (Fig. 1g). In vertebrates, Tsg has a dual function, binding to both BMP and Chordin to form a ternary complex that can either promote or inhibit BMP signalling depending on the level of Tolloid[12,13]. Thus, in the early amphioxus gastrula, *Tsg* may primarily function in inhibiting BMP signalling dorsally. Later, *Tsg* expression extends more ventrally, but never overlaps with that of *Tolloid-like*. Because the functions of BMPs, Chordin, Tsg and Tolloid proteins are evolutionarily conserved[8], our results indicate that, as in vertebrates, the interaction of these proteins may establish a BMP-signalling gradient in the amphioxus gastrula that is high ventrally and low dorsally.

The secreted proteins Nodal and its antagonist, Lefty, are both produced by the gastrula organizer in vertebrates[8–10]. Similarly, in the amphioxus early–mid-gastrula, *Nodal*[14] and *Lefty* are expressed dorsally in both mesendoderm and ectoderm (Fig. 2a, b). Expression of both genes is downregulated posteriorly by the late gastrula and on the right side by the early neurula (Fig. 2a, b), reminiscent of their asymmetric expression in vertebrate embryos[15]. These results indicate that amphioxus, like vertebrates, may use Nodal signalling for both dorsoventral and left/right patterning.

Expression of organizer-related transcription factors

Several genes encoding transcription factors, including *even-skipped* (*Evx*), *Hex*, *Vent* and *Goosecoid* mediate signalling of secreted patterning proteins in the vertebrate gastrula. All of their amphioxus orthologues are expressed in patterns comparable to those of their vertebrate counterparts. For example, in both amphioxus and vertebrates[16], *Evx* is expressed ventroposteriorly[17] (Fig. 2c). Similarly, *Hex*, an anterior endodermal marker in vertebrates[10,18], is expressed in presumptive endoderm at the centre of the vegetal plate in amphioxus (Fig. 2d). *Hex* is always expressed more anteriorly and ventrally than *Chordin* (Figs 1c and 2d, h), indicating that in amphioxus, as proposed for vertebrates[1,19], *Hex*-expressing anterior endoderm may represent a separate signalling centre for A/P patterning of the gastrula. *Vent* is expressed ventrolaterally in early embryos of both amphioxus[20] (Fig. 2e) and vertebrates, and represses expression of *Chordin*[9]. Double *in situ* hybridization shows that in amphioxus expression of *Vent* and *Chordin* is complementary (Fig. 2e, g). *Goosecoid* expression is also similar in amphioxus and *Xenopus laevis*, being expressed at the dorsal lip of the blastopore of the early gastrula in both organisms[21] (Fig. 2f). However, by the mid-gastrula stage there are minor differences, with *Goosecoid* expression migrating anteriorly to the prechordal mesendoderm in *Xenopus*[1], but extending throughout the axial mesoderm in amphioxus (Fig. 2f).

BMP protein treatment ventralizes amphioxus embryos

To test whether BMP signalling functions in D/V patterning of the amphioxus embryo as it does in vertebrates[8,10], we added zebrafish recombinant BMP4 protein (zBMP4) to cultures of amphioxus blastulae to globally elevate signalling, as done previously in ascidian and hemichordate embryos[22,23]. Amphioxus embryos are ventralized in a dose-dependent manner. We observed more severe reduction of dorsal structures with increasing zBMP concentrations (data not shown). At 250 ng ml^{-1}, zBMP4 protein caused a uniform phenotype (Fig. 3a–d) similar to that resulting from overexpression of BMP2 or BMP4 in vertebrate embryos[24,25]. Amphioxus embryos treated with zBMP4 are shortened and lack dorsal structures, including the notochord, neural tube and somites. However, the anterior tip of the embryo and the tail fin remain recognizable, suggesting that the anterior and posterior ends of the embryo were properly specified.

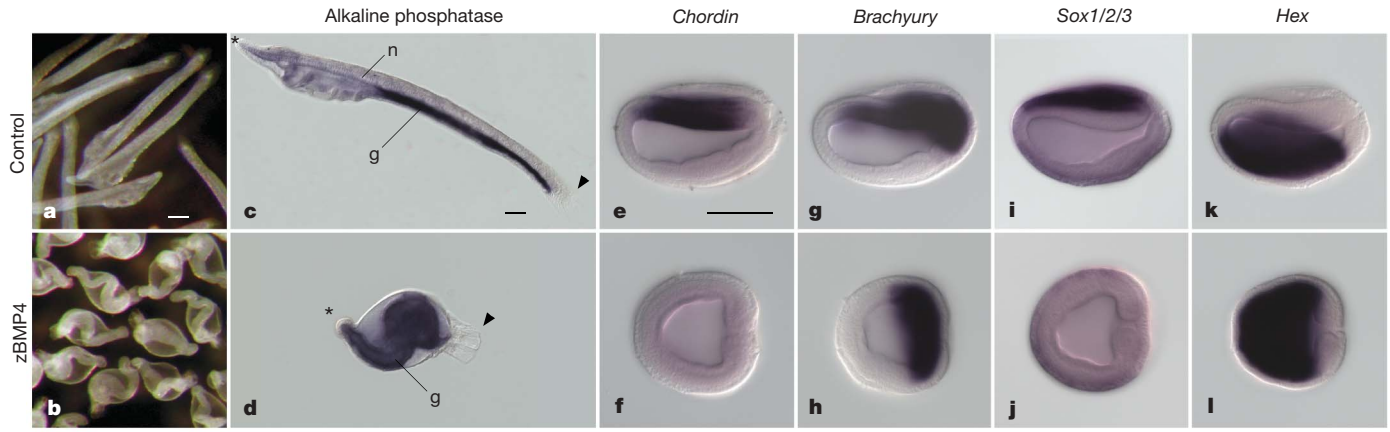


Figure 3 | Exogenous zebrafish BMP4 protein (zBMP4) ventralizes amphioxus embryos. Upper panels (**a**, **c**, **e**, **g**, **i**, **k**) show controls cultured with 250 ng ml^{-1} bovine serum albumin. Lower panels (**b**, **d**, **f**, **h**, **j**, **l**) show embryos treated with 250 ng ml^{-1} zBMP4. **a**, **b**, Embryos at 36 h. zBMP4 application results in a uniform phenotype. Scale bar, 100 μm. **c**, **d**, Staining for alkaline phosphatase, a marker of mid- and hindgut endoderm (g), in control and in zBMP4 treated embryos at 36 h post-fertilization. Asterisk, anterior tip of the embryo. Arrowheads, tail fin; n, notochord. Scale, 50 μm. **e**–**l**, expression of D/V markers in control and zBMP-treated early neurulae. All *in situ* images are side views; anterior towards the left. Scale bar, 50 μm.

Staining for the endodermal marker—endogenous alkaline phosphatase—showed that endoderm is present, but entirely specified as mid- and hindgut, with no pharynx (Fig. 3c, d). The expression of dorsal markers confirmed ventralization of the embryos. In zBMP-treated embryos at the onset of neurulation, the *Chordin* domain in dorsal axial mesoderm (notochord) is eliminated and that of *Brachyury* expression is considerably reduced (Fig. 3e, f). However, the *Brachyury* domain in the circumblastoporal mesendoderm is unaffected (Fig. 3g, h). Expression of *Sox1/2/3*, a neural plate marker, is also eliminated, suggesting absence of the central nervous system (Fig. 3i, j). In contrast, *Hex* expression, normally restricted to ventral endoderm, is expanded throughout the entire mesendoderm (Fig. 3k, l). Thus, exogenous BMP4 protein treatment repressed expression of markers of the dorsal mesoderm and central nervous system and

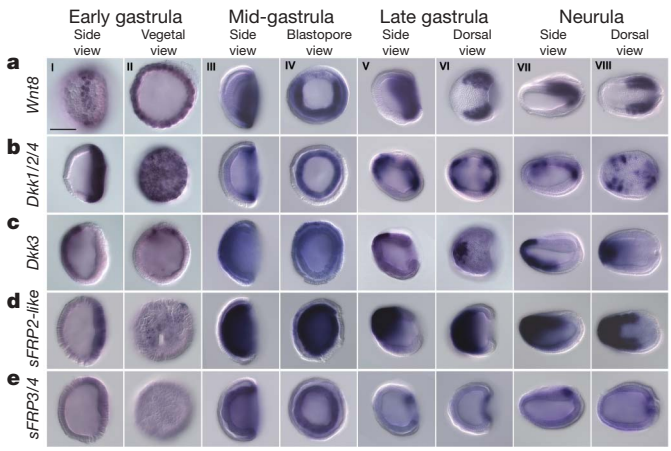


Figure 4 | Expression of amphioxus *Wnt8* and *Wnt* antagonists during embryogenesis. Anterior towards the left in side and dorsal views. **a**, *Wnt8* is expressed posteroventrally around the blastopore. **b**, *Dkk1/2/4* is expressed in the invaginating vegetal plate and later in anterior mesendoderm and around the blastopore. **c**, *Dkk3* is expressed in anterior ectoderm and mesendoderm. **d**, *sFRP2-like* is first expressed in the vegetal plate and later in the anterior-half of the embryo. **e**, *sFRP3/4* is initially expressed throughout the mesendoderm, after the late gastrula stage expression is restricted to the blastopore. Expression of *sFRP3/4* is undetectable at the early gastrula stage. Scale bar, 50 μm.

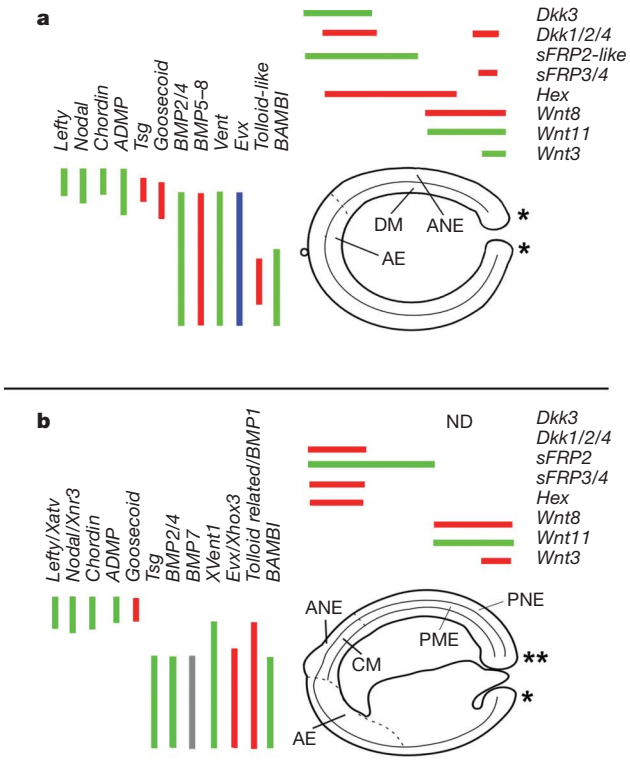


Figure 5 | Schematic diagram of the expression of dorso/ventral- and anterior/posterior-patterning genes at the late gastrula stage. **a**, Amphioxus. **b**, Frog, stage 12.5. Anterior at left; blastopore at right. Expression of D/V patterning genes (irrespective of A/P distribution) at left. Expression of A/P patterning genes (irrespective of D/V distribution) at right. Single asterisk, little involution over blastopore lip. Green bars, expression in both mesendoderm and ectoderm; red bars, expression only in mesendoderm; blue bar, expression in ectoderm; grey bar, tissue layers expressed in have not been determined. Double asterisk, extensive involution over blastopore lip. AE, anterior endoderm; ANE, anterior neuroectoderm; CM, chordamesoderm; DM, dorsal mesoderm; ND, not determined; PME, posterior mesendoderm; PNE, posterior neuroectoderm.

expanded ventral endoderm markers, suggesting that BMP signalling has an important role in amphioxus D/V patterning.

Expression of Wnt signalling molecules

In addition to BMP and Nodal proteins, Wnts are important secreted signalling molecules in early development and are involved in both D/V and A/P patterning. Accordingly, the organizer is a source of many Wnt antagonists[8,10]. To examine conservation of Wnt signals across chordates, we determined the expression of several genes involved in Wnt signalling during amphioxus gastrulation. Expression of *Wnt*s was generally restricted to the posterior half of the amphioxus gastrula, suggesting a posterior Wnt-signalling centre. *Wnt8* is the first amphioxus *Wnt* to be zygotically expressed[26]. At the onset of gastrulation, it is expressed throughout the vegetal plate but soon becomes restricted to a ring at the ectoderm/mesendoderm boundary (Fig. 4a) where it is co-expressed with *Brachyury*[6]. By the mid-gastrula, *Wnt8* is downregulated dorsally (Fig. 4a), consistent with an additional role in D/V patterning. At this stage, transcripts of several other Wnts and nuclear β-catenin are also localized around the blastopore[26,27] (Fig. 5; Supplementary Fig. 11). To determine if, as in vertebrates, Wnt signalling is likely to be suppressed dorsally and anteriorly in amphioxus, we also examined expression of several Wnt antagonists (*Dkk1/2/4*, *Dkk3*, *sFRP2-like* and *sFRP3/4*). Their expression during gastrulation differs somewhat from that of their vertebrate orthologues. Amphioxus *Dkk3* is strongly expressed in the anterior ectoderm at the onset of gastrulation and by mid-gastrula is also weakly expressed in the underlying mesendoderm. Later, expression becomes restricted to anterior endoderm and anterior/dorsal ectoderm (Fig. 4c). Expression of vertebrate *Dkk3*, which may suppress Wnt signalling[28] although it is generally not thought to be a Wnt antagonist, has not been described at the gastrula stage. The other three vertebrate genes are expressed in the organizer in the early gastrula, and by the late gastrula, are all expressed in the prechordal mesoderm, although *sFRP2* is also expressed in the prospective neuroectoderm[8,29]. In amphioxus, by the late gastrula, expression of *Dkk1/2/4* and *sFRP2-like* resembles that of their vertebrate homologues, although that of *sFRP3/4* is different. *Dkk1/2/4*, which is homologous to three vertebrate genes, initially switches on in the amphioxus vegetal plate concomitant with the downregulation of *Wnt8*. As gastrulation proceeds, both genes remain expressed in mesendoderm around the blastopore, but at the late gastrula, *Dkk1/2/4* also turns on in the anterior mesendoderm (Fig. 4b). The two amphioxus *sFRP* genes are congruently expressed throughout the mesendoderm at the mid-gastrula, but by the late gastrula, their domains are complementary: *sFRP2-like* becomes restricted to anterior ectoderm and mesendoderm and *sFRP3/4* to the rim of the blastopore (Fig. 4d, e). Thus, expression of amphioxus Wnt antagonists is limited to anterior and posterior portions of the embryo, consistent with a role for Wnt/β-catenin signalling in A/P patterning.

Discussion

Our data indicate that the deployment of early embryonic patterning molecules along the D/V and A/P axes of amphioxus embryos is much like that in vertebrates and demonstrate a conserved role for BMP signalling in D/V patterning. Co-expression of amphioxus orthologues of many vertebrate organizer genes (for example, *Chordin*, *ADMP*, *Nodal*, *Lefty*, *Goosecoid* and *Lim1/5*; ref. 30) at the dorsal blastopore lip, together with previous results from grafting experiments[7], suggest that a dorsal organizer using BMP antagonists and Nodal and its antagonist Lefty to mediate D/V patterning was present before the split of cephalochordates and vertebrates. Recent phylogenetic analyses with large gene sets have reversed the positions of amphioxus and tunicates, placing amphioxus basally within the chordates and tunicates as the sister group of vertebrates. Thus, although it was suggested on the basis of differences in expression of organizer genes in tunicates[31,32] and vertebrates that the organizer may have arisen after the split of the tunicates from the amphioxus plus vertebrate lineage[2], it seems likely that the organizer was present at the base of the chordates, but perhaps lost in tunicates.

At least some elements of these conserved D/V mechanisms are present in more basal deuterostomes. In the direct-developing hemichordate *Saccoglossus kowalevskii*, several BMPs and BMP antagonists are also expressed on opposite sides of the embryo but in an inverted D/V orientation compared with amphioxus and vertebrates[23]. BMP/Chordin antagonism also functions in D/V patterning in protostomes such as *Drosophila*, with *Chordin* and *BMP* expressed as in hemichordates on the ventral and dorsal sides respectively. Together with our data, these comparisons support the hypothesis that the inversion of the D/V axis[8] occurred at the base of the chordates[23]. Our data also suggest that the phylogenetic grouping of amphioxus with echinoderms proposed by Delsuc *et al.*[33] is improbable as it would require either separate inversion of the D/V axis in the amphioxus and vertebrate lineages or an inversion at the base of deuterostomes with a subsequent reversion in the hemichordates. Thus, our findings support the canonical chordate grouping as including cephalochordates, tunicates and vertebrates, as suggested by a recent analysis of deuterostome phylogeny[5].

There is no evidence that a role of BMP/Chordin antagonism in D/V patterning arose before the bilaterians. Although homologues of *BMP*s and *Chordin* are expressed asymmetrically in early embryos of the cnidarian *Nematostella vectensis*, they do not seem to function in D/V patterning, but rather in segregating the germ layers[34]. It has been proposed that the deployment of BMP antagonists during gastrulation for D/V patterning might have been recruited with the evolution of the definitive mesoderm in triploblastic animals[34,35]. Thus, it seems that the role of BMP/Chordin antagonism in D/V patterning evolved in bilaterians before the protostome/deuterostome split, whereas the inversion of the D/V axis[8] occurred at the base of the chordates.

Our data also indicate that the early amphioxus embryo may have separate signalling centres to establish gradients of BMP signalling along the D/V axis and Wnt signalling along the A/P axis. These findings provide important insights into the evolutionary origin of the signalling centres that constitute the vertebrate organizer functions. Many amphioxus *Wnt*s are co-expressed with *Brachyury* posteriorly around the blastopore, and their orthologues are also expressed at the posterior pole of protostomes and at the mouth/anus end of cnidarians, suggesting that a blastoporal Wnt-signalling centre predated the split of cnidarians and bilaterians[36–38]. Expression domains of amphioxus Wnt antagonists are predominately anterior and have limited overlap with those of the BMP antagonist *Chordin*, whereas those of *Hex* and *Chordin* fail to overlap at all. Such separate signalling centres are reminiscent of those in chick and mouse[10,19] in which an anterior signalling centre in the *Hex*- and *Dkk1*-expressing anterior hypoblast (chick) or anterior visceral endoderm (mouse) is involved in head development and is distinct from the node or organizer. In contrast, in the early *Xenopus* gastrula, the *Hex/Dkk1*-expressing anterior endoderm along with head and trunk organizers are all located in close proximity to the dorsal blastopore lip[10]. However, separate clusters of genes respond to overexpression of Wnt and BMP antagonists at early gastrula stages, indicating that head and trunk organizers involved, respectively, in patterning the A/P and D/V axes may be functionally distinct at an early stage[39], subsequently becoming spatially separated as involution carries the anterior endoderm and head organizer forward. Because involution is all but absent during amphioxus gastrulation[6], it is likely that axial patterning in amphioxus, with separate signalling centres producing BMP and Wnt signalling gradients along the D/V and A/P axes, respectively, reflects the ancestral state of axial patterning for chordates.

METHODS

Obtaining animals and embryos. Sexually mature adults of the Florida amphioxus (*Branchiostoma floridae*) were collected by shovel and sieve in Tampa Bay,

Florida, during the summer breeding season. Gametes from electrically stimulated adults were fertilized, and the embryos and larvae were raised in the laboratory.

Library construction and EST sequencing. Five libraries were used in this study—from unfertilized eggs, gastrula, neurula, 36-h larvae and mature adults. Total RNA was isolated by acid guanidium isothiocyanate–phenol–chloroform method, and poly(A)$^+$ RNA with a Dynabeads messenger RNA purification kit (Invitrogen, Carlsbad California). The CloneMiner cDNA library construction kit (Invitrogen, Carlsbad California) was used to generate the five libraries. Each of the libraries was arrayed in 384-well plates by a Genetix Q-Pix robot (Genetix, Boston, Massachusetts). ESTs of nearly 140,000 cDNA clones from the five libraries were sequenced at the Academia DNA Sequencing Center, National Institute of Genetics, Mishima, Japan. As clones were grouped into clusters using 3′ ESTs, all clones in each cluster encode the same gene.

Identification of organizer-related genes. Human, mouse, chick, frog and zebrafish protein sequences for organizer genes were used as queries to tBLASTn search our amphioxus EST database. We retrieved hit EST sequences with *E* values $< 10^{-5}$ and performed BLASTx searches against the International Protein Index proteome set released on July 12, 2005. When the best-hit sequence of the EST corresponded to the starting vertebrate protein query, the cDNA cluster was assigned as the putative orthologue, and the longest clone from each cluster was picked. We subsequently sequenced the complete cDNA insert by primer walking and did molecular phylogenetic analysis to confirm the orthology assignment (see Supplementary Methods and Supplementary Figs 1–10 for details). Further details for *in situ* hybridization of organizer-related genes are provided in the Supplementary Methods section.

Culture of embryos with zBMP protein. Cultures of amphioxus were treated with recombinant zebrafish BMP4 (zBMP4, R&D systems) at concentrations from 10 ng ml^{-1} to 250 ng ml^{-1}. As for control experiments, the same concentrations of bovine serum albumin were applied separately to the same stage of cultures. Proteins were added from the early blastula stage (2.5 h post-fertilization) and embryos were fixed for alkaline phosphatase staining or *in situ* hybridization at different time points during development.

Received 29 May; accepted 20 November 2006.
Published online 21 January 2007.

1. Gerhart, J. Evolution of the organizer and the chordate body plan. *Int. J. Dev. Biol.* **45,** 133–153 (2001).
2. Kourakis, M. J. & Smith, W. C. Did the first chordates organize without the organizer? *Trends Genet.* **21,** 506–510 (2005).
3. Blair, J. E. & Hedges, S. B. Molecular phylogeny and divergence times of deuterostome animals. *Mol. Biol. Evol.* **22,** 2275–2284 (2005).
4. Philippe, H., Lartillot, N. & Brinkmann, H. Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia. *Mol. Biol. Evol.* **22,** 1246–1253 (2005).
5. Bourlat, S. J. *et al.* Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* **444,** 85–88 (2006).
6. Zhang, S.-C., Holland, N. D. & Holland, L. Z. Topographic changes in nascent and early mesoderm in amphioxus embryos studied by DiI labeling and by *in situ* hybridization for a *Brachyury* gene. *Dev. Genes Evol.* **206,** 532–535 (1997).
7. Tung, T. C., Wu, S. C. & Tung, Y. Y. F. Experimental studies on the neural induction in amphioxus. *Sci. Sin.* **11,** 805–820 (1962).
8. De Robertis, E. M. & Kuroda, H. Dorsal–ventral patterning and neural induction in *Xenopus* embryos. *Annu. Rev. Cell Dev. Biol.* **20,** 285–308 (2004).
9. De Robertis, E. M., Larrain, J., Oelgeschlager, M. & Wessely, O. The establishment of Spemann's organizer and patterning of the vertebrate embryo. *Nature Rev. Genet.* **1,** 171–181 (2000).
10. Niehrs, C. Regionally specific induction by the Spemann–Mangold organizer. *Nature Rev. Genet.* **5,** 425–434 (2004).
11. Reversade, B. & De Robertis, E. M. Regulation of ADMP and BMP2/4/7 at opposite embryonic poles generates a self-regulating morphogenetic field. *Cell* **123,** 1147–1160 (2005).
12. Chang, C. *et al.* Twisted gastrulation can function as a BMP antagonist. *Nature* **410,** 483–487 (2001).
13. Scott, I. C. *et al.* Homologues of Twisted gastrulation are extracellular cofactors in antagonism of BMP signalling. *Nature* **410,** 475–478 (2001).
14. Yu, J. K., Holland, L. Z. & Holland, N. D. An amphioxus *nodal* gene (*AmphiNodal*) with early symmetrical expression in the organizer and mesoderm and later asymmetrical expression associated with left-right axis formation. *Evol. Dev.* **4,** 418–425 (2002).
15. Levin, M. Left–right asymmetry in embryonic development: a comprehensive review. *Mech. Dev.* **122,** 3–25 (2005).
16. Joly, J. S., Joly, C., Schulte-Merker, S., Boulekbache, H. & Condamine, H. The ventral and posterior expression of the zebrafish homeobox gene *eve1* is perturbed in dorsalized and mutant embryos. *Development* **119,** 1261–1275 (1993).
17. Ferrier, D. E., Minguillon, C., Cebrian, C. & Garcia-Fernandez, J. Amphioxus *Evx* genes: implications for the evolution of the midbrain–hindbrain boundary and the chordate tailbud. *Dev. Biol.* **237,** 270–281 (2001).
18. Jones, C. M., Broadbent, J., Thomas, P. Q., Smith, J. C. & Beddington, R. S. An anterior signalling centre in *Xenopus* revealed by the homeobox gene *XHex*. *Curr. Biol.* **9,** 946–954 (1999).
19. Stern, C. D. Initial patterning of the central nervous system: how many organizers? *Nature Rev. Neurosci.* **2,** 92–98 (2001).
20. Kozmik, Z. *et al.* Characterization of amphioxus *amphivent*, an evolutionarily conserved marker for chordate ventral mesoderm. *Genesis* **29,** 172–179 (2001).
21. Neidert, A. H., Panopoulou, G. & Langeland, J. A. Amphioxus *goosecoid* and the evolution of the head organizer and prechordal plate. *Evol. Dev.* **2,** 303–310 (2000).
22. Darras, S. & Nishida, H. The BMP/CHORDIN antagonism controls sensory pigment cell specification and differentiation in the ascidian embryo. *Dev. Biol.* **236,** 271–288 (2001).
23. Lowe, C. J. *et al.* Dorsoventral patterning in hemichordates: insights into early chordate evolution. *PLoS Biol.* **4,** e291 (2006).
24. Dale, L., Howes, G., Price, B. M. & Smith, J. C. Bone morphogenetic protein 4: a ventralizing factor in early *Xenopus* development. *Development* **115,** 573–585 (1992).
25. Nikaido, M., Tada, M., Saji, T. & Ueno, N. Conservation of BMP signaling in zebrafish mesoderm patterning. *Mech. Dev.* **61,** 75–88 (1997).
26. Schubert, M. & Holland, L. Z. in *Wnt Signaling in Development* (ed. Kuhl, M.) 210–239 (Eurakah, 2003).
27. Holland, L. Z., Panfilio, K. A., Chastain, R., Schubert, M. & Holland, N. D. Nuclear β-catenin promotes non-neural ectoderm and posterior cell fates in amphioxus embryos. *Dev. Dyn.* **233,** 1430–1443 (2005).
28. Hoang, B. H. *et al.* Dickkopf 3 inhibits invasion and motility of Saos-2 osteosarcoma cells by modulating the Wnt-β-catenin pathway. *Cancer Res.* **64,** 2734–2739 (2004).
29. Pera, E. M. & De Robertis, E. M. A direct screen for secreted proteins in *Xenopus* embryos identifies distinct activities for the Wnt antagonists Crescent and Frzb-1. *Mech. Dev.* **96,** 183–195 (2000).
30. Langeland, J. A., Holland, L. Z., Chastain, R. A. & Holland, N. D. An amphioxus LIM-homeobox gene, *AmphiLim1/5*, expressed early in the invaginating organizer region and later in differentiating cells of the kidney and central nervous system. *Int. J. Biol. Sci.* **2,** 110–116 (2006).
31. Imai, K. S., Levine, M., Satoh, N. & Satou, Y. Regulatory blueprint for a chordate embryo. *Science* **312,** 1183–1187 (2006).
32. Passamaneck, Y. J. & Di Gregorio, A. *Ciona intestinalis*: chordate development made simple. *Dev. Dyn.* **233,** 1–19 (2005).
33. Delsuc, F., Brinkmann, H., Chourrout, D. & Philippe, H. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **439,** 965–968 (2006).
34. Matus, D. Q., Thomsen, G. H. & Martindale, M. Q. Dorso/ventral genes are asymmetrically expressed and involved in germ-layer demarcation during cnidarian gastrulation. *Curr. Biol.* **16,** 499–505 (2006).
35. Matus, D. Q. *et al.* Molecular evidence for deep evolutionary roots of bilaterality in animal development. *Proc. Natl Acad. Sci. USA* **103,** 11195–11200 (2006).
36. Broun, M., Gee, L., Reinhardt, B. & Bode, H. R. Formation of the head organizer in hydra involves the canonical Wnt pathway. *Development* **132,** 2907–2916 (2005).
37. Holland, L. Z. Heads or tails? Amphioxus and the evolution of anterior–posterior patterning in deuterostomes. *Dev. Biol.* **241,** 209–228 (2002).
38. Kusserow, A. *et al.* Unexpected complexity of the *Wnt* gene family in a sea anemone. *Nature* **433,** 156–160 (2005).
39. Hufton, A. L., Vinayagam, A., Suhai, S. & Baker, J. C. Genomic analysis of *Xenopus* organizer function. *BMC Dev. Biol.* **6,** 27, doi: 10.1186/1471-213X-6-27 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are indebted to J.M. Lawrence, University of South Florida, for providing laboratory facilities during the summer breeding season of amphioxus. This work was funded by grants from the National Science Foundation, USA (L.Z.H. and N.D.H.), the National Aeronautics and Space Administration, USA (M.B.-F. and L.Z.H.), the National Institutes of Health (M.B.-F.), MEXT, Japan (N.S. and Y.K.), and the 21st Century COE for the Biodiversity Research at Kyoto University (N.S.). J.-K.Y is currently supported by the Della Martin prize postdoctoral fellowship from the Division of Biology, California Institute of Technology, USA.

Author Information EST sequences were deposited in the DDBJ/EMBL/GenBank databases. (DNA Data Bank of Japan accession numbers BW692960–BW954996). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to L.Z.H. (lzholland@ucsd.edu).

ARTICLES

Quantitative dynamics and binding studies of the 20S proteasome by NMR

Remco Sprangers[1] & Lewis E. Kay[1]

The machinery used by the cell to perform essential biological processes is made up of large molecular assemblies. One such complex, the proteasome, is the central molecular machine for removal of damaged and misfolded proteins from the cell. Here we show that for the 670-kilodalton 20S proteasome core particle it is possible to overcome the molecular weight limitations that have traditionally hampered quantitative nuclear magnetic resonance (NMR) spectroscopy studies of such large systems. This is achieved by using an isotope labelling scheme where isoleucine, leucine and valine methyls are protonated in an otherwise highly deuterated background in concert with experiments that preserve the lifetimes of the resulting NMR signals. The methodology has been applied to the 20S core particle to reveal functionally important motions and interactions by recording spectra on complexes with molecular weights of up to a megadalton. Our results establish that NMR spectroscopy can provide detailed insight into supra-molecular structures over an order of magnitude larger than those routinely studied using methodology that is generally applicable.

The proteasome plays a critical role in the degradation of proteins that regulate crucial cellular processes such as cell division, gene expression, signal transduction and apoptosis and is a target for anti-cancer drugs[1–3]. The 670-kDa 20S core particle (CP) consists of four heptameric rings arranged in an $\alpha_7\beta_7\beta_7\alpha_7$ fashion[4] (Fig. 1a). The two outer α_7 rings form the entrance channel for substrates and the binding sites for the poly-ubiquitin-recognizing 19S proteasome[5] and the immune-response-related 11S activator[6]. The interfaces between the α_7 and β_7 rings enclose the two antechambers, where substrates can be stored before degradation[7]. The two inner β_7 rings encompass the catalytic chamber, where the active sites are sequestered within the lumen of the barrel. Although high-resolution structures of the 20S proteasome from all three domains of life (eukarya, bacteria, archaea) are available[4,8–10], little direct information on the internal motions of the CP can be obtained from these static pictures. However, dynamics are likely to be critical for the function of this enzyme, especially for the α-domains that gate the entrance of targets for proteolysis[11].

Methyl transverse relaxation optimized spectroscopy (TROSY)

In general, the structural features of many stable complexes can be determined by electron microscopy and X-ray crystallography, but insight into functionally important motions and transient interactions are more difficult to obtain. NMR spectroscopy is especially suited to the study of dynamics and dynamic interactions over a broad spectrum of timescales; however, applications have been limited to relatively small proteins[12]. The study of megadalton-sized complexes by solution NMR spectroscopy is challenging and requires careful consideration of what probes to use (isotope labelling), along with optimization of experiments that specifically account for the spin properties of these probes[12]. We have recently shown that very-high-quality ^1H,^{13}C methyl spectra can be obtained for highly deuterated proteins that are protonated at isoleucine-δ1, valine and leucine methyl positions (Fig. 1b–e, Fig. 2a, b)[13].

Here we have used this labelling scheme, along with methyl-TROSY spectroscopy that optimizes both sensitivity and resolution in spectra[14]. Figure 1b shows the heteronuclear multiple quantum coherence (HMQC) spectrum of the isoleucine-δ1 methyl region of the archaeabacterium *Thermoplasma acidophilum* CP, which was recorded on a 25-μM sample (0.35-mM monomer concentration) at 65 °C in 90 min on an 800-MHz spectrometer with a room-temperature probehead. With the exception of one cross-peak, all of the expected 17 (α-ring) and 14 (β-ring) correlations from isoleucine residues are observed in the spectrum. There are a number of practical advantages in studying the *T. acidophilum* CP. First, it is structurally very similar to eukaryotic versions[2]. Second, all of the α-subunits are equivalent in this archaeal CP (as are the β-subunits), leading to improvements in both sensitivity and resolution in NMR spectra. Third, its thermal stability allows studies at elevated temperatures, lowering the molecular tumbling time τ_C and thus improving spectral quality ($\tau_C \approx 180$ ns at 65 °C; by comparison for a 15-kDa protein at 30 °C, a τ_C value of ~8 ns would be expected). Fourth, this 20S CP can be reconstituted from separately expressed subunits and it is therefore possible to restrict isotopic labelling to either of the α- or β-rings, making the other subunits invisible to NMR.

Methyl chemical shift assignments

Here we report the nearly complete chemical shift assignment for isoleucine-δ1, leucine and valine methyl groups in the α-subunits of the 670-kDa CP, despite the fact that backbone-directed triple-resonance NMR methods fail. The strategy used to solve this assignment problem involves 'dissecting' the large system into smaller building blocks (Supplementary Fig. 1). The chemical shifts of the residues in the core of these building blocks will not be affected much by the disruption of the complex provided that the smaller units fold in a similar manner in isolation and in the complex.

Here, the 670-kDa 20S proteasome CP (Fig. 1a–c) was dissected into the 360-kDa double-α-ring particle ($\alpha_7\alpha_7$, Fig. 1a, d) that spontaneously forms from wild-type (WT) α monomers in the absence of β subunits[15]. Subsequently, a monomeric form of the α protein ($\alpha_{monomer}$; Fig. 1a, e) was generated by introducing mutations that prevent formation of $\alpha_7\alpha_7$ (see Methods). This 21-kDa $\alpha_{monomer}$ was assigned using standard triple-resonance TROSY-based NMR experiments[16,17]. Both secondary chemical shifts and nuclear

[1]Departments of Biochemistry, Medical Genetics and Chemistry, The University of Toronto, Toronto, Ontario M5S 1A8, Canada.

Overhauser effect (NOE) cross-peaks confirm that the structure of the $\alpha_{monomer}$ is similar to that of the α-subunit in the crystal structure of the full 20S CP, and indeed, spectra of all three forms of the complex (Fig. 1a) show similarities (Fig. 1c–e, Supplementary Fig. 2). Despite the large molecular weight of the 360-kDa $\alpha_7\alpha_7$ complex, correlation maps linking methyl groups with the carbons one and two bonds removed could be recorded using methyl-TROSY versions of previously developed correlated spectroscopy (COSY) experiments[18] (Fig. 2a). Four chemical shifts ($^1H_{methyl}$, $^{13}C_{methyl}$ and two for $^{13}C_{side\text{-}chain}$) could therefore be compared between $\alpha_{monomer}$ and $\alpha_7\alpha_7$; where the correspondence is high, assignments of methyl groups could be transferred from $\alpha_{monomer}$ to $\alpha_7\alpha_7$ (Fig. 2a) and confirmed on the basis of NOE patterns (Fig. 2b).

In the next step (Supplementary Fig. 1), additional assignments were obtained for methyl groups proximal to those already assigned, using NOE correlations and the high-resolution crystal structure[4]. Successively, methyl assignments from $\alpha_7\alpha_7$ were transferred to the full 670-kDa $\alpha_7\beta_7\beta_7\alpha_7$ proteasome, based on methyl chemical shifts (Fig. 1c, d), NOE patterns between residues (Fig. 2b) and NOE cross-peaks within residues (Fig. 2c). Finally, resonances for which no NOE or COSY correlations were observed were assigned via mutagenesis[19]. In our hands this approach is only successful once the majority of assignments are available, because mutations can have substantial effects on the chemical shifts of neighbouring methyl groups (Supplementary Fig. 3). The combined analysis of the data produced assignments for 95% (92 out of 97) and 89% (86 out of 97) of the isoleucine, leucine, and valine methyl groups of $\alpha_7\alpha_7$ and the CP, respectively. Reporters for 96% ($\alpha_7\alpha_7$) and 91% ($\alpha_7\beta_7\beta_7\alpha_7$) of these methyl-containing residues are available for the study of site-specific dynamics and interactions.

Probing molecular dynamics

Insights into pico- to nanosecond timescale side-chain dynamics in the proteasome were obtained by recording the relaxation properties of both ^2H and ^{13}C spins[20,21] in $^{13}CHD_2$-labelled samples. Figure 3a shows the correlation between ^2H- and ^{13}C-derived order parameters squared S^2 for the $\alpha_7\alpha_7$ complex, which quantify the amplitude of motion (with the amplitude decreasing as S^2 increases). We observe an excellent correlation between the two independent measurements, despite the size of the complex (360 kDa, 50 °C, $\tau_C \approx 120$ s), so that the measures of the dynamics can be interpreted with confidence (Supplementary Fig. 4). Values of S^2 were also obtained for the 670-kDa CP (65 °C, $\tau_C \approx 180$ s) and show a very strong correlation with those in $\alpha_7\alpha_7$ (Supplementary Fig. 5). Thus, the β-rings have little effect on either the structure or the dynamics of the α-rings, apart from residues at the α,β interface that are less flexible in the context of the β-rings (for example, V107, V113). We note that although most residues showing increased temperature factors[4,22] do experience fast timescale mobility, there is little correlation between crystallographic B-factors and S^2 values (Supplementary Fig. 6).

Plotting S^2 on the structure of the proteasome reveals a cluster of flexible residues on the outside of the barrel, between individual α subunits (Fig. 3b). Interestingly, α-subunits of the CP from *Mycobacterial tuberculosis* and *Rhodococcus* as well as the eukaryotic α6 subunits show sequence deletions in part of this region. This highly mobile cluster carries a potential nuclear localization-type signal. Although the archaea-bacterium does not possess a nucleus, the flexible signal sequence from the *T. acidophilum* CP has been shown to promote nuclear localization in human cell lines[23] and is conserved in the α2-subunit of the mammalian proteasome[9]. It may be that the mobile sequence was present before nuclei evolved and was the target for different adaptor proteins early in evolution[23]. Irrespective of the identities of the binding partners the dynamics at this site probably play a part in moderating the affinity of interactions by providing an entropic barrier to binding.

Micro- to millisecond timescale dynamics are often correlated with protein function so we studied these motions in the proteasome

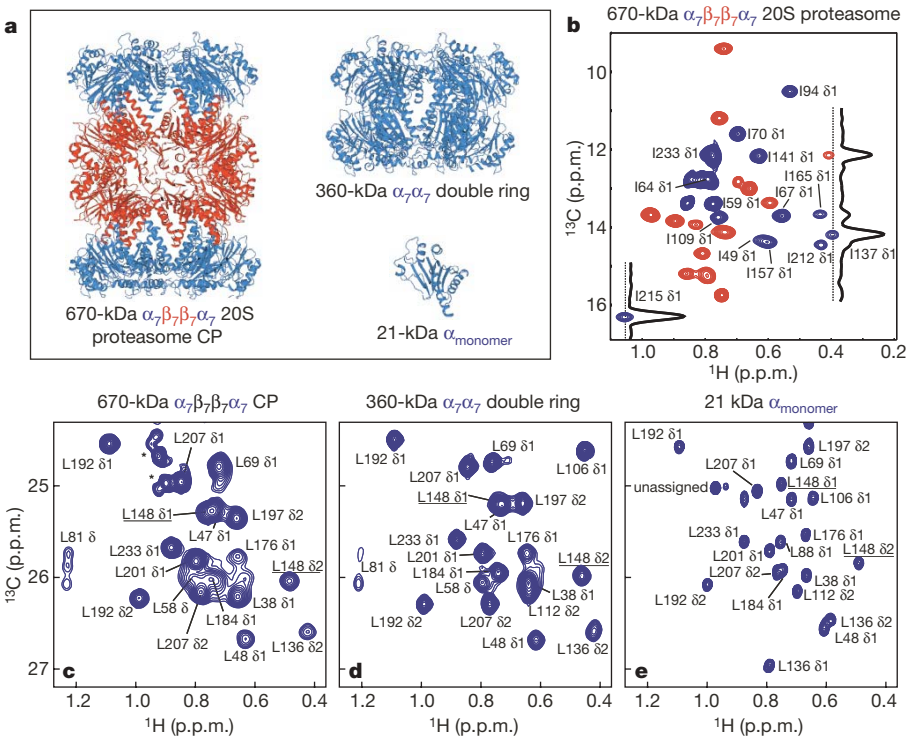


Figure 1 | Details of the proteasome structure and methyl assignments. **a**, Ribbon diagrams of the 20S CP ($\alpha_7\beta_7\beta_7\alpha_7$), the $\alpha_7\alpha_7$ double ring and the $\alpha_{monomer}$ (mutated as described in Methods). **b**–**e**, Methyl-TROSY spectra (800 MHz) of the CP (65 °C), $\alpha_7\alpha_7$ and $\alpha_{monomer}$ (50 °C). Correlations belonging to residues within α or β subunits are colour-coded blue or red, respectively. We note that, apart from the spectrum in **b**, only the α-subunit is NMR-active. The position of L148, one of the starting points in the assignment of the proteasome, is indicated by underlining (see Fig. 2). The correlations in **c**, indicated by asterisks, are not present in all preparations and probably derive from impurities. The complete methyl-TROSY spectra are shown in Supplementary Fig. 2.

using relaxation dispersion experiments. ^{1}H-^{13}C multiple-quantum methyl-TROSY dispersion profiles[24] (Supplementary Fig. 7) measured for both $\alpha_7\alpha_7$ and the CP fitted well to the simplest model of chemical exchange ($A \underset{k_B}{\overset{k_A}{\rightleftarrows}} B$) and similar exchange values ($k_{ex} = k_A + k_B$) were obtained for all residues, which suggests that the source of the exchange involves a single global process. Values of k_{ex} of 1,200 ± 170 and 1,580 ± 90 s^{-1} were obtained for $\alpha_7\alpha_7$ and the CP, respectively, with the small differences probably reflecting the different temperatures at which the experiments were performed (50 °C versus 65 °C). Interestingly, the exchanging residues cluster on the inside of the antechamber and form a surface from the entrance pore of the proteasome towards the catalytic chamber (Fig. 3c); this highly dynamic environment could indirectly facilitate the movement of substrate to the sites of proteolysis by presenting an entropic barrier to binding to the walls of both the antechamber and the entrance pores that would prevent the deposit of protein in these regions. Whether the correlated motions in the antechamber are specific for the proteasome or are more widespread in oligomeric molecular machines has still to be determined; however, studies of many protein complexes are now within the range of these NMR approaches.

In addition to residues inside the antechamber, the amino-terminal residues of the α subunits are mobile on the millisecond timescale (V14 is extremely broadened and no resonance could be assigned to I12, with spectra of WT and I12A CP showing no changes in the isoleucine-δ1 region). The 12 N-terminal residues are not observed in the crystal structure of the archaeal proteasome[4], but the corresponding residues in the yeast and bovine proteasomes[8,9] form a stable conformation that closes the entrance channel and functions as a gate[11]. NMR dispersion data on a mutant protein that lacks residues 1 to 12 (the Δ1–12 CP) establish that these residues have little effect on the observed dynamics inside the antechambers. However, a comparison of HMQC spectra of the WT and Δ1–12 CP show chemical shift changes at the narrowest point of the substrate entrance path (V129; Fig. 3b, c) and inside the antechamber that suggest that the 12 N-terminal residues fold into the antechamber through the entrance channel to form a primitive gate (Fig. 3d). This observation is consistent with electron microscopy data that shows density for the residues around the entrance pore[25], but goes beyond the available structural data by showing that at least some of these dynamic termini probably fold into the antechamber. It also explains the fact that these 'invisible' gating residues influence peptide degradation rates[25]. Interestingly, some of the N-terminal residues of the eukaryotic proteasome also fold back into the lumen of the antechamber[8,9] and thereby block substrate entrance[6,11].

Quantitative studies of binding

Protein machines do not function in isolation and it is important to understand the quaternary organization of these large complexes. NMR is a particularly powerful tool in this regard because many interactions in transient complexes are too weak to be detected and characterized using other structural techniques. Unfortunately, the size limitations associated with traditional NMR experiments have limited progress in the study of large complexes. Recently, however, ^{1}H-^{15}N cross-correlated relaxation-induced polarization transfer (CRIPT) spectroscopy was used to establish the binding site on the 70-kDa GroES of the 800-kDa GroEL[26].

Here we used methyl assignments to determine the residues of the 670-kDa CP that interact with the 11S activator complex from *Trypanosoma brucei* in the 1.1 MDa 11S-$\alpha_7\beta_7\beta_7\alpha_7$-11S complex[6]. Stepwise addition of the 11S complex to the CP leads to the disappearance of a subset of resonances and to the appearance of a new set of resonances (Fig. 4a). Residues affected by binding are

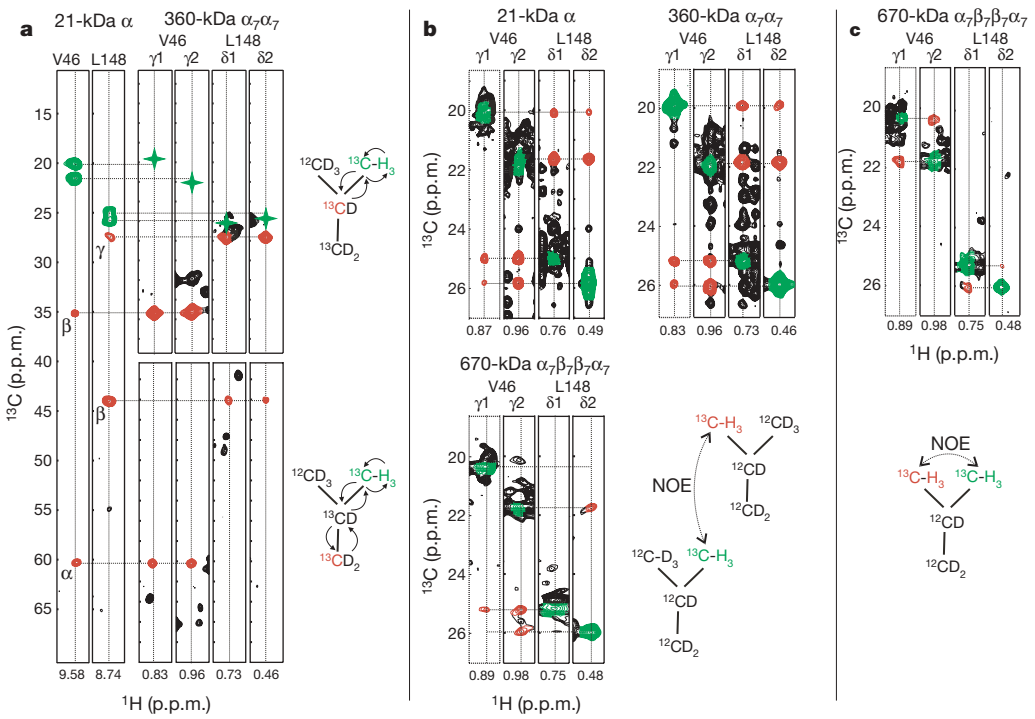


Figure 2 | Details of the assignment strategy. a, Side-chain carbon chemical shifts obtained from an (H)C(CO)NH TOCSY (total correlation spectroscopy) experiment of the $\alpha_{monomer}$ (left two strips) correlate well with the side-chain chemical shifts in $\alpha_7\alpha_7$ obtained from out-and-back $H_{methyl}C_{methyl}C_{\beta/\gamma}$ and $H_{methyl}C_{methyl}(C_{\beta/\gamma})C_{\alpha/\beta}$ methyl-TROSY, COSY-based experiments[18] (right four strips). The green stars in the COSY strips indicate the positions of the diagonal peaks that are either not observed or are very weak in the experiment. The labelling pattern and magnetization transfer for the methyl-detected experiments are indicated. **b**, NOESY strips showing correlations between proximal methyls of V46 and L148, along with the labelling pattern and magnetization transfer (during the 250-ms mixing time). We note that there are no intra-residue NOE cross-peaks, because only one of the two methyl groups is labelled for each residue. **c**, NOESY strips from an experiment recorded on the $\alpha_7\beta_7\beta_7\alpha_7$ CP (50-ms mixing time) where both methyls of leucine and valine residues in the α-rings are labelled as indicated, giving cross-peaks between methyl groups of the same residue.

located in the interaction site that has been identified by crystallography (Fig. 4b, c), proving that despite the megadalton size of the 11S activator-proteasome complex, binding interfaces can be accurately identified. The high quality of the data can be appreciated from the fact that a microscopic dissociation constant (Fig. 4d and Supplementary Information) was obtained for the $\alpha_7\alpha_7$/11S interaction from fits of decreases in peak intensities of the unbound form and concomitant increases in correlations reporting on the bound state ($12 \pm 10\,\mu M$); titration data recorded for the full 20S proteasome (at a concentration of only $9\,\mu M$) are consistent with a similar affinity interaction (inset to Fig. 4d). Although the 20S CP/11S complex is non-cognate, the structural features of the interaction are conserved[22], and it has been shown that the 11S from *Trypanosoma brucei* activates both rat and yeast proteasomes, in addition to the 20S archaeal CP studied here[6,22].

In summary, this study establishes that quantitative information of the sort normally reserved for applications to small proteins in solution can be obtained on systems more than an order of magnitude larger than those usually studied by NMR spectroscopy. A combined approach in which X-ray crystallography is used to produce detailed static pictures, followed by NMR studies of motion and ligand interactions promises to be particularly fruitful for the study of supramolecular structures. In particular, by exploiting the methodological advances described here, we have identified and quantified specific sites of motion that may facilitate substrate localization to the catalytic chamber, and potentially modulate interactions of the nuclear localization-type signal with receptors. We have also presented data that is consistent with the N termini of the α-subunits that make up the dynamic gate folding through the entrance pore into the antechamber. In addition, the 11S binding site has been mapped and its affinity measured. Future quantitative studies of the interaction of the proteasome with dynamical substrates, such as α-synuclein[27], and with substrates inside the lumen of the protease[7] are now feasible, and will lead to additional insights into the function of the proteasome[1].

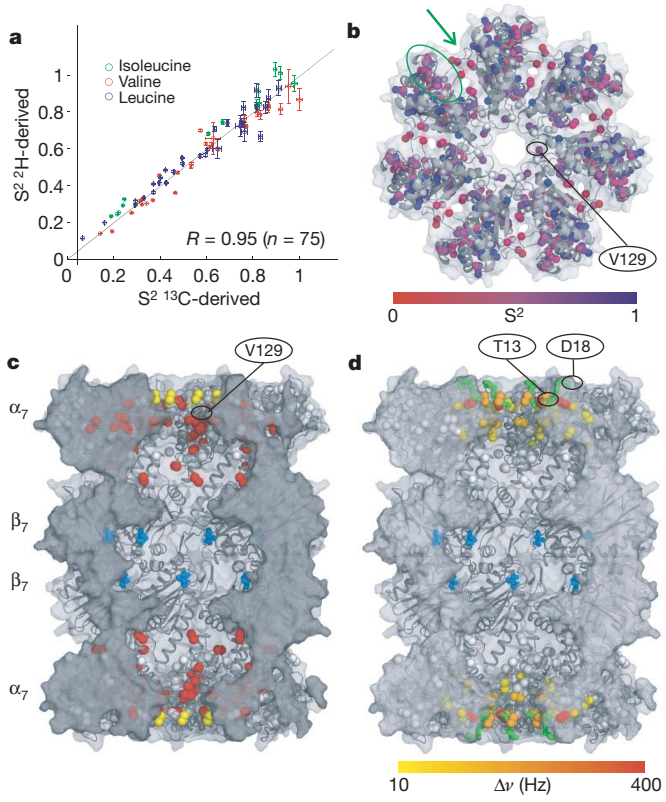


Figure 3 | Quantification of dynamics and structure. **a**, Correlation between S^2 values extracted independently from carbon[21] and deuterium[20] relaxation experiments. Errors (±1 standard deviation) are indicated with bars[21]. **b**, Top view of the proteasome with methyl groups coloured according to S^2. A cluster of very flexible side chains is located between the individual monomers (circle, arrow), with the variable loop containing the potential nuclear localization-type signal indicated by an arrow (see text). **c**, Cross-section of the side view of the proteasome that reveals the inside (lumen). Residues undergoing concerted motion are shown in red and form a continuous cluster in the antechamber. The exchange-broadened V14 in the flexible N terminus is shown in yellow, while the active site threonines are indicated in blue. **d**, Side view highlighting residues that change chemical shift upon truncation of the first 12 residues of the α subunit (chemical shift differences $\Delta\nu$ between WT and Δ1–12, defined as $(\Delta\nu_H^2 + \Delta\nu_C^2)^{1/2}$ at 800 MHz ^{1}H resonance frequency). The location of residues 13–18 in the crystal structure (green) is shown and suggests that the invisible N-terminal 12 residues[4] are located in the lumen of the antechamber. Although the chemical shift differences were measured on $\alpha_7\alpha_7$, for clarity the complete proteasome is shown.

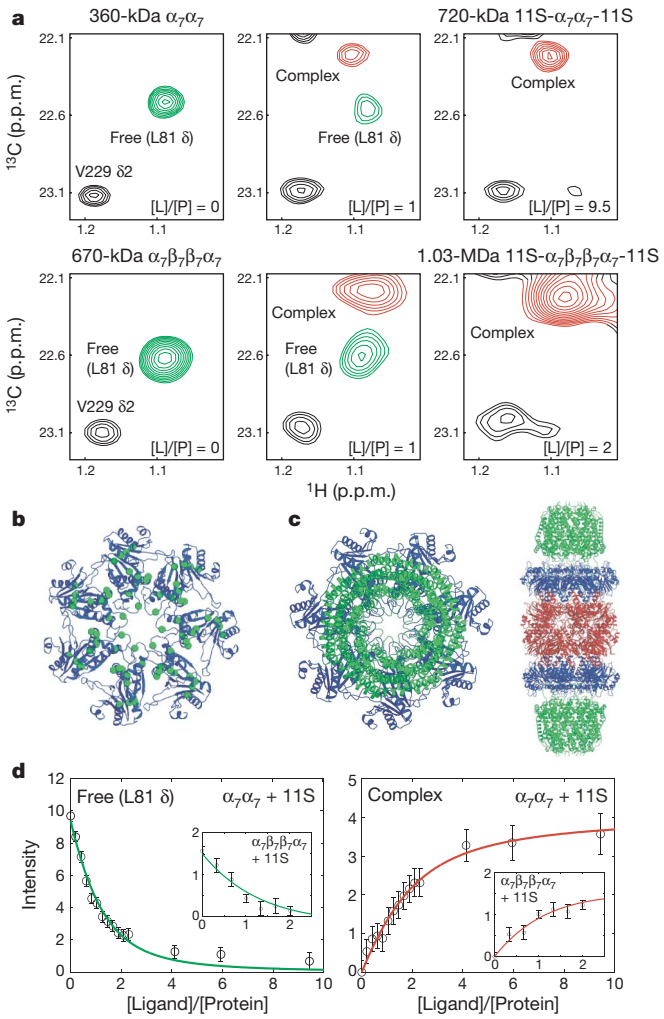


Figure 4 | Interaction between the 20S proteasome and the 11S activator. **a**, Methyl-TROSY spectra of $\alpha_7\alpha_7$ (top) and the CP (bottom) during addition of the 11S activator complex. The complex is in slow exchange on the NMR chemical shift timescale; assignments of correlations from the complex are not available. **b**, Residues whose resonances are affected by the 11S interaction are mapped on the proteasome structure. **c**, Crystal structure of the 11S-proteasome complex[6,22]. **d**, The intensities of resonances during the titration are used to obtain an approximate dissociation constant ($K_D = 12 \pm 10\,\mu M$) for the $\alpha_7\alpha_7$/11S interaction that is consistent with the CP titration data (see insets). The decrease in intensity of one of the correlations from L81 is shown in green and the concomitant increase in a 'bound' peak indicated in red. Errors are quantified from signal-to-noise in spectra. [Ligand] and [Protein] refer to total ligand and protein concentrations.

METHODS

Sample preparation. Proteins were produced using labelled precursors and growth media as described in the Supplementary Information. All proteins were purified on Ni-NTA resin (Qiagen), followed by cleavage of the purification tag using tobacco etch virus (TEV) protease and gel filtration. The full proteasome was obtained from reconstitution of separately purified NMR-active α-subunits and NMR-inactive β-subunits or from co-expression of both subunits. Proteasome assembly was verified by electron microscopy (data not shown). Final protein concentrations of the NMR samples were between 0.1 and 2.8 mM in monomer (see Supplementary Information) in 100% D_2O (or 90% H_2O/10% D_2O), 25 mM potassium phosphate, pH 6.8, 50 mM NaCl, 1 mM EDTA, 0.03% NaN_3 and 2 mM DTT.

NMR spectroscopy. NMR spectra (see also Supplementary Information) were recorded at 50 °C ($\alpha_{monomer}$, $\alpha_7\alpha_7$) or 65 °C (CP = $\alpha_7\beta_7\beta_7\alpha_7$) using 600 and 800 MHz Varian Inova spectrometers equipped with cryogenically cooled (600 MHz) or room-temperature (800 MHz) pulsed-field gradient triple-resonance probes; all data sets of the CP were measured at 800 MHz. ^{1}H and ^{13}C chemical shift assignments are referenced against DSS (2,2-dimethyl-2-silapentane-5-sulphonic acid). Backbone and side-chain assignments of $\alpha_{monomer}$ (Δ2–34, R57A, R86A, R130A) were achieved using standard TROSY-based triple-resonance methods[16,17]. Stereo-specific assignments were made using a 10% ^{13}C labelled protein[28] and, where possible, transferred to $\alpha_7\alpha_7$ and the CP on the basis of peak positions and relative intensities in NOE spectra. Methyl-TROSY[14] COSY-based out-and-back data sets were measured on the $\alpha_7\alpha_7$ sample using pulse schemes that are very similar to those previously published[18], with the exception of the methyl ^{13}C evolution period, which was of the HMQC variety. All NOE spectra were recorded using methyl-TROSY-based schemes with ^{13}C chemical shifts obtained in (t_1,t_2) and with mixing times as described in Supplementary Information. Details of relaxation experiments are given in Supplementary Information.

Chemical-shift titration experiments were recorded on a 14 μM $\alpha_7\alpha_7$ complex (0.2 mM monomer concentration) and on a 9 μM CP complex (0.13 mM monomer concentration) by stepwise addition of 0.2 mM 11S activator complex (1.4 mM monomer concentration) and K_D values extracted as described in Supplementary information. All data were processed with the NmrPipe/NmrDraw suite of programs[29]. Figures displaying molecular structures were made with PyMol (www.pymol.sourceforge.net).

Received 28 July; accepted 6 December 2006.
Published online 21 January 2007.

1. Pickart, C. M. & Cohen, R. E. Proteasomes and their kin: proteases in the machine age. *Nature Rev. Mol. Cell Biol.* **5,** 177–187 (2004).
2. Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92,** 367–380 (1998).
3. Adams, J. The proteasome: a suitable antineoplastic target. *Nature Rev. Cancer* **4,** 349–360 (2004).
4. Löwe, J. *et al.* Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science* **268,** 533–539 (1995).
5. Glickman, M. H. *et al.* A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. *Cell* **94,** 615–623 (1998).
6. Whitby, F. G. *et al.* Structural basis for the activation of 20S proteasomes by 11S regulators. *Nature* **408,** 115–120 (2000).
7. Sharon, M. *et al.* 20S proteasomes have the potential to keep substrates in store for continual degradation. *J. Biol. Chem.* **281,** 9569–9575 (2006).
8. Groll, M. *et al.* Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* **386,** 463–471 (1997).
9. Unno, M. *et al.* The structure of the mammalian 20S proteasome at 2.75 Å resolution. *Structure* **10,** 609–618 (2002).
10. Kwon, Y. D., Nagy, I., Adams, P. D., Baumeister, W. & Jap, B. K. Crystal structures of the *Rhodococcus* proteasome with and without its pro-peptides: implications for the role of the pro-peptide in proteasome assembly. *J. Mol. Biol.* **335,** 233–245 (2004).
11. Groll, M. *et al.* A gated channel into the proteasome core particle. *Nature Struct. Biol.* **7,** 1062–1067 (2000).
12. Mittermaier, A. & Kay, L. E. New tools provide new insights in NMR studies of protein dynamics. *Science* **312,** 224–228 (2006).
13. Tugarinov, V., Hwang, P. M. & Kay, L. E. Nuclear magnetic resonance spectroscopy of high-molecular-weight proteins. *Annu. Rev. Biochem.* **73,** 107–146 (2004).
14. Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E. & Kay, L. E. Cross-correlated relaxation enhanced ^{1}H-^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.* **125,** 10420–10428 (2003).
15. Zwickl, P., Kleinz, J. & Baumeister, W. Critical elements in proteasome assembly. *Nature Struct. Biol.* **1,** 765–770 (1994).
16. Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl Acad. Sci. USA* **94,** 12366–12371 (1997).
17. Salzmann, M., Pervushin, K., Wider, G., Senn, H. & Wüthrich, K. TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc. Natl Acad. Sci. USA* **95,** 13585–13590 (1998).
18. Tugarinov, V. & Kay, L. E. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J. Am. Chem. Soc.* **125,** 13868–13878 (2003).
19. Sprangers, R., Gribun, A., Hwang, P. M., Houry, W. A. & Kay, L. E. Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. *Proc. Natl Acad. Sci. USA* **102,** 16678–16683 (2005).
20. Tugarinov, V., Ollerenshaw, J. E. & Kay, L. E. Probing side-chain dynamics in high molecular weight proteins by deuterium NMR spin relaxation: an application to an 82-kDa enzyme. *J. Am. Chem. Soc.* **127,** 8214–8225 (2005).
21. Tugarinov, V. & Kay, L. E. Quantitative ^{13}C and ^{2}H NMR relaxation studies of the 723-residue enzyme malate synthase G reveal a dynamic binding interface. *Biochemistry* **44,** 15970–15977 (2005).
22. Forster, A., Masters, E. I., Whitby, F. G., Robinson, H. & Hill, C. P. The 1.9 Å structure of a proteasome-11S activator complex and implications for proteasome-PAN/PA700 interactions. *Mol. Cell* **18,** 589–599 (2005).
23. Nederlof, P. M., Wang, H. R. & Baumeister, W. Nuclear localization signals of human and *Thermoplasma* proteasomal alpha subunits are functional *in vitro*. *Proc. Natl Acad. Sci. USA* **92,** 12060–12064 (1995).
24. Korzhnev, D. M., Kloiber, K., Kanelis, V., Tugarinov, V. & Kay, L. E. Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: application to a 723-residue enzyme. *J. Am. Chem. Soc.* **126,** 3964–3973 (2004).
25. Benaroudj, N., Zwickl, P., Seemuller, E., Baumeister, W. & Goldberg, A. L. ATP hydrolysis by the proteasome regulatory complex PAN serves multiple functions in protein degradation. *Mol. Cell* **11,** 69–78 (2003).
26. Fiaux, J., Bertelsen, E. B., Horwich, A. L. & Wüthrich, K. NMR analysis of a 900K GroEL GroES complex. *Nature* **418,** 207–211 (2002).
27. Liu, C. W., Corboy, M. J., DeMartino, G. N. & Thomas, P. J. Endoproteolytic activity of the proteasome. *Science* **299,** 408–411 (2003).
28. Neri, D., Szyperski, T., Otting, G., Senn, H. & Wüthrich, K. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry* **28,** 7510–7516 (1989).
29. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6,** 277–293 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Forman-Kay for discussions and for providing laboratory space, F. Hansen for discussions, J. Rubenstein for electron microscopy images, R. Muhandiram for NMR support and C. Hill for a plasmid of the 11S activator complex. R.S. acknowledges EMBO and the Canadian Institutes of Health Research (CIHR) Training Program in Protein Folding and Disease for fellowships. L.E.K. holds a Canada Research Chair in Biochemistry. Grant support from CIHR and NSERC is acknowledged.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to L.E.K. (kay@pound.med.utoronto.ca).

LETTERS

Coherent control of optical information with matter wave dynamics

Naomi S. Ginsberg[1], Sean R. Garner[1] & Lene Vestergaard Hau[1]

In recent years, significant progress has been achieved in manipulating matter with light, and light with matter[1]. Resonant laser fields interacting with cold, dense atom clouds provide a particularly rich system[2–6]. Such light fields interact strongly with the internal electrons of the atoms, and couple directly to external atomic motion through recoil momenta imparted when photons are absorbed and emitted. Ultraslow light propagation in Bose–Einstein condensates[7] represents an extreme example of resonant light manipulation using cold atoms. Here we demonstrate that a slow light pulse can be stopped and stored in one Bose–Einstein condensate and subsequently revived from a totally different condensate, 160 μm away; information is transferred through conversion of the optical pulse into a travelling matter wave. In the presence of an optical coupling field, a probe laser pulse is first injected into one of the condensates where it is spatially compressed to a length much shorter than the coherent extent of the condensate. The coupling field is then turned off, leaving the atoms in the first condensate in quantum superposition states that comprise a stationary component and a recoiling component in a different internal state. The amplitude and phase of the spatially localized light pulse are imprinted on the recoiling part of the wavefunction, which moves towards the second condensate. When this 'messenger' atom pulse is embedded in the second condensate, the system is re-illuminated with the coupling laser. The probe light is driven back on and the messenger pulse is coherently added to the matter field of the second condensate by way of slow-light-mediated atomic matter-wave amplification. The revived light pulse records the relative amplitude and phase between the recoiling atomic imprint and the revival condensate. Our results provide a dramatic demonstration of coherent optical information processing with matter wave dynamics. Such quantum control may find application in quantum information processing and wavefunction sculpting.

In our experiments, two Bose–Einstein condensates[8–10] (BECs) of approximately 1.8×10^6 sodium atoms each, and separated by more than 100 μm, are created in a double-well trapping potential in internal energy state $|1\rangle$(Fig. 1a). The trapping potential is abruptly turned off, and after 1 ms, the atoms are illuminated with a 'coupling' laser beam resonant with their internal $|2\rangle \rightarrow |3\rangle$ transition and travelling in the $-z$ direction (Fig. 1a, b). A counter-propagating, 3 μs gaussian 'probe' laser pulse, resonant with the $|1\rangle \rightarrow |3\rangle$ transition, is then injected into the first BEC. The laser beams drive the atoms into coherent superposition 'dark' states with destructively interfering absorption amplitudes such that neither probe nor coupling laser is absorbed[11–13]. The propagating light pulse creates a slight atomic polarization that slows and spatially compresses the pulse by a factor of 5×10^7 (refs 7, 14, 15). Ultimately, the light pulse is completely contained within the condensate[16].

With $\psi_i(\mathbf{R}, t)$, $(i = 1, 2, 3)$ representing the three components of an atom's external wavefunction at time t and position $\mathbf{R}$, the dark state superposition is $\Psi_D(\mathbf{R}, t) = \psi_1(\mathbf{R}, t)|1\rangle + \psi_2(\mathbf{R}, t)|2\rangle$, where the amplitude and phase of the state $|2\rangle$ component relative to the state $|1\rangle$ component are determined by the amplitude and phase of the probe electric field, $\mathbf{E}_p(\mathbf{R}, t) = \frac{1}{2}\boldsymbol{\mathcal{E}}_p(\mathbf{R}, t)e^{i(\mathbf{k}_p\cdot\mathbf{R}-\omega_p t)} + \text{c.c.}$ (where c.c. is the complex conjugate), relative to the coupling electric field, $\mathbf{E}_c(\mathbf{R}, t) = \frac{1}{2}\boldsymbol{\mathcal{E}}_c(\mathbf{R}, t)e^{i(\mathbf{k}_c\cdot\mathbf{R}-\omega_c t)} + \text{c.c.}$, according to[17]:

$$\frac{\psi_2(\mathbf{R}, t)}{\psi_1(\mathbf{R}, t)} = -\frac{\Omega_p(\mathbf{R}, t)}{\Omega_c(\mathbf{R}, t)}e^{i(\mathbf{k}_p-\mathbf{k}_c)\cdot\mathbf{R}-i(\omega_p-\omega_c)t} \quad (1)$$

Here, $\Omega_p(\mathbf{R}, t) = \mathbf{d}_{31}\cdot\boldsymbol{\mathcal{E}}_p(\mathbf{R}, t)/\hbar$ and $\Omega_c(\mathbf{R}, t) = \mathbf{d}_{32}\cdot\boldsymbol{\mathcal{E}}_c(\mathbf{R}, t)/\hbar$ are the probe and coupling Rabi frequencies, where $\mathbf{d}_{jk} = -e\langle j|\mathbf{r}|k\rangle$ are electric dipole matrix elements, $\boldsymbol{\mathcal{E}}_{p,c}$ are the slowly varying envelopes of the laser fields, $-e$ is the electron charge, and $\hbar$ is Planck's constant. In the present experiment, the anti-parallel orientation of the probe and coupling wavevectors, $\mathbf{k}_p$ and $\mathbf{k}_c$, produces phase variation in the dark state on optical length scales.

According to equation (1), the variation of ψ_2 in space and time mimics that of the probe light pulse such that a slowly varying envelope of ψ_2 accompanies that of the highly compressed, slowly moving light pulse through the condensate. With the light pulse thus contained in the condensate, we turn the coupling laser off over 40 ns. As a result, in order to preserve the dark state (equation (1)), the atoms coherently and adiabatically drive the probe light field to extinction, but the dark state imprint of the pulse remains in the atom cloud[16,17,18]. The spatial phase variation impressed on ψ_2 in this process corresponds to a two-photon recoil of $\hbar(k_p + k_c)/m = 59\,\mu\text{m}\,\text{ms}^{-1}$, where m is the atomic mass. Hence a ψ_2 'messenger' atom pulse is ejected from its initial position[19] in the same direction as the incident light pulse, and ultimately leaves the first condensate and travels as a coherent matter wave towards, through and beyond the second BEC (Fig. 1c).

When this messenger pulse is embedded in the second BEC (for example, 2.1 ms in Fig. 1c), we re-illuminate the system with the coupling laser. Even though the messenger pulse is alien to this second BEC, the atoms cooperatively drive the light pulse back on. The revived light pulse then propagates out of the second condensate under slow light conditions, with the propagation direction $(+z)$ determined by the phase imprinted on the messenger atoms[17]. A light pulse revived after 2.7 ms of messenger flight is detected on a photomultiplier tube, as presented in Fig. 2a.

A comparison of Figs 2a and 3a reveals that a pulse revived in a separate BEC appears qualitatively similar to one revived inside the condensate in which it was stored. Before discussing the physics of light pulse storage and revival in separate BECs, we examine the process for a single atom cloud. Light pulses can be revived in a single BEC up to 0.7 ms after storage, that is, as long as the messenger atom

[1]Department of Physics, and Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA.

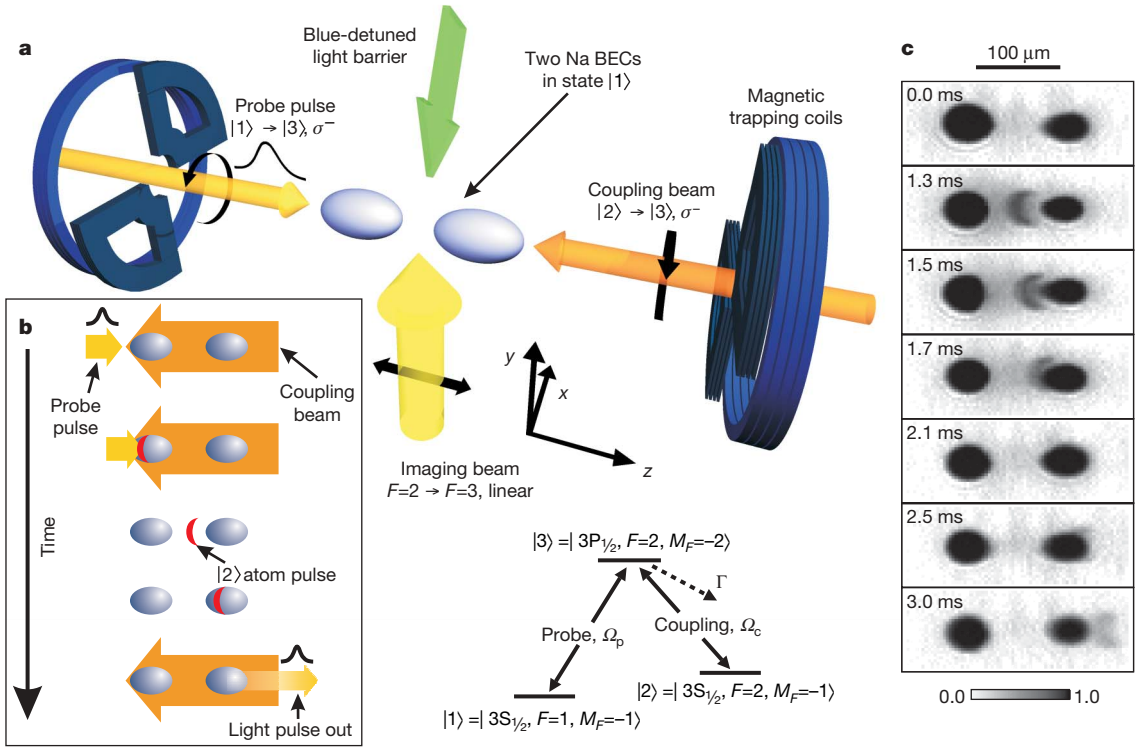


Figure 1 | Diagram of the experiment. **a**, Two sodium BECs (pale blue) are prepared in state $|1\rangle$ in a double-well potential formed by combining a harmonic magnetic trap (dark blue) and a repulsive optical dipole barrier (focused 532 nm green laser beam with elliptical gaussian cross-section). The entire potential is turned off 1 ms before experiments begin, whereupon the probe (light orange) and coupling (orange) laser beams are introduced. Finally, the condensates are imaged with a laser beam (yellow), near resonance for the atoms' $F = 2 \rightarrow F = 3$ transition, after optical pumping to $F = 2$. **b**, Experiments begin with the injection of a probe laser pulse (light orange) into the first (left) BEC, while the cloud is illuminated by the counter-propagating coupling laser beam (orange). The pulse propagates into the condensate under ultraslow light conditions. After the probe pulse is spatially compressed within the cloud, the coupling beam is switched off, leaving an imprint of the probe pulse's phase and amplitude in the form of atomic population amplitude in state $|2\rangle$ (red). Each atom's $|2\rangle$ component has a momentum corresponding to two photon recoils (absorption from the probe beam and stimulated emission into the coupling beam) and is ejected towards the second (right) BEC. When this 'messenger' atom pulse in $|2\rangle$ arrives, the coupling beam is switched back on, and the probe light pulse is regenerated in the second condensate. Revived light pulses are imaged onto a 50 μm pinhole (to reject background light) and detected with a photomultiplier tube. **c**, Resonant absorption images of BECs and travelling messenger pulse, at indicated times since light pulse storage. No revival coupling beam is fired, and the messenger pulse is observed to travel through and beyond the second BEC.

pulse remains within the condensate. In this case, we can describe the revival of a light pulse as resulting from an interference of each atom's wavefunction with itself. The ψ_2 imprint is much smaller than the extent of the ψ_1 wavefunction and has been translated (owing to two-photon recoil) from its original location (see Fig. 3a inset). During revival, the coupling laser creates some ψ_3 amplitude sourced from the translated ψ_2 imprint, and hence a dipole moment, proportional to $\psi_1^*(\mathbf{R})\psi_3(\mathbf{R})$, is generated in each atom, which drives the probe light field back on. As the system is driven into the dark state, described by equation (1), the ratio between the value of the initial ψ_1 at the storage location and of ψ_1 at the revival position is mapped onto the regenerated light pulse. The ψ_2 imprint is coherently added to ψ_1 at the revival location as the light pulse subsequently leaves the region under slow light conditions. This picture shows that a direct measurement of the one-body density matrix of the atom cloud[20] can be made by recording the revived light pulse energy as a function of the distance between the storage and revival locations. That light pulse revival is possible over the full length of a Bose condensate (Fig. 3a) reflects the condensate's off-diagonal long-range order[21]. In contrast, in the non-condensed atom cloud shown in Fig. 3b, the revived pulse energy decays with a 1/e time of 2.5 μs during which the $|2\rangle$ component travels only 147 nm. This distance is of the order of the thermal de Broglie wavelength and is much less than the extent of the cloud.

When light pulse storage and revival occur in two different atom clouds separated before condensation, as in Fig. 2a, each atom's wavefunction is initially localized to either but not both of the two isolated BECs. Therefore, the dark state superposition imprinted during storage exists only for atoms from the first BEC. Nevertheless, a coherent light pulse can still be revived from the second condensate through bosonic matter wave stimulation. To see how, we use a second quantized description of the matter fields. The interaction between light and matter is governed by the hamiltonian

$$\hat{\mathbf{H}}_{\mathrm{int}} = -\frac{1}{2}\int \mathrm{d}\mathbf{R}(\mathbf{d}_{31}^{*}\cdot\hat{\mathbf{E}}_{\mathrm{p}}^{(-)}(\mathbf{R})\hat{\psi}_1^{\dagger}(\mathbf{R})\hat{\psi}_3(\mathbf{R}) + \mathbf{d}_{32}\cdot\hat{\mathbf{E}}_{\mathrm{c}}^{(+)}(\mathbf{R})\hat{\psi}_3^{\dagger}(\mathbf{R})\hat{\psi}_2(\mathbf{R}) + \mathrm{h.c.}) \quad (2)$$

where $\hat{\psi}_i^{\dagger}(\mathbf{R})$ and $\hat{\psi}_i(\mathbf{R})$ are creation and annihilation operators for an atom in internal state $|i\rangle$ at position $\mathbf{R}$ (ref. 22), and h.c. indicates the hermitian conjugate. Here, we have also expressed the laser fields in second quantized form[23] to stress the symmetry of the matter and light fields discussed below. During revival, when the coupling laser creates a population amplitude in $|3\rangle$ (second term in equation (2)), the presence of a BEC in $|1\rangle$ creates a large rate for bosonic stimulated scattering of atoms into the condensate mode (due to the presence of $\hat{\psi}_1^{\dagger}$ in the first term of equation (2)[24–28]). This bosonic stimulation drives the probe light field on, and the four interaction terms of $\hat{\mathbf{H}}_{\mathrm{int}}$ in combination drive the system into a dark state[17]. In this picture, the coupling laser field and the matter field for atoms in $|1\rangle$ form a perfectly symmetric pair: bosonic stimulation into a macroscopically occupied photon field (the coupling laser) drives coherent dynamics

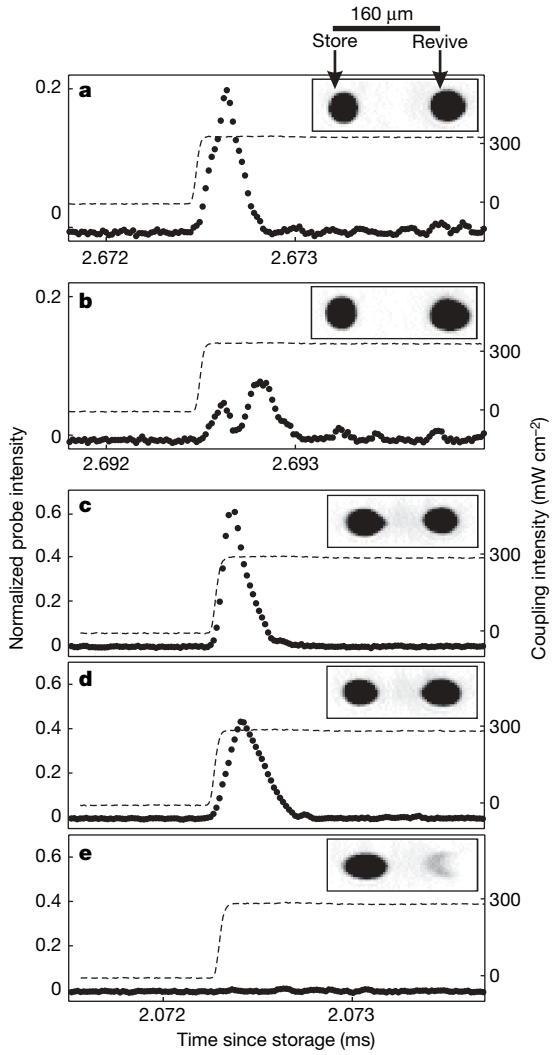


Figure 2 | Light pulse storage and revival in two separate condensates. Revived probe pulses, normalized to input pulse intensity, are plotted against time since pulse storage (dots, left-hand axis), and simultaneously recorded coupling intensity (dashed line, right-hand axis). Insets are resonant absorption images of the $|1\rangle$ condensates, 20 μs after revival. **a**, **b**, Light pulses revived in the second of a pair of independently condensed BECs. Atoms are evaporatively cooled in a 2.3-μK-deep double-well potential formed by a magnetic trap combined with a light barrier (the Bose-condensation temperature is 660 nK). After condensation, the magnetic potential is adiabatically softened to $\omega_z = 2\pi \times 20$ Hz and $\omega_r = 2\pi \times 40$ Hz. Subsequently, the light barrier is adiabatically lowered to 10μ, where μ is each well's resulting chemical potential. The trapping potential is then turned off in less than 200 μs. After 1 ms, the probe pulse is stored in the first BEC ($\Omega_p = 2\pi \times 2.6$ MHz, $\Omega_c = 2\pi \times 2.6$ MHz). In **a**, the light pulse is revived in the second BEC after 2.67 ms during which time the $|2\rangle$ atom pulse travels 157 μm. Note, $\Omega_{c,revival} = 2\pi \times 21.4$ MHz, resulting in a temporally narrowed output pulse[16]. In **b**, the messenger $|2\rangle$ pulse travels to a different location in the second BEC, where differences in density and phase patterns between the two lead to a bimodal structure. **c**, **d**, Revived light pulses from condensates formed by adiabatically splitting a single magnetically trapped BEC with a 1.5μ-tall light barrier, ramped up over 100 ms, and held constant for 1 s. A typical pulse (**c**) contains 6.9×10^3 photons, 2.2% of the input pulse energy. In **d**, a larger, denser second BEC yields a slower light propagation speed and a broader and less intense pulse, with similar energy to **c**. **e**, A control experiment in which experimental timing and conditions were exactly the same as in **c** and **d**, but with no second BEC.

during the initial light pulse injection, whereas stimulation into a macroscopically occupied matter field (the second $|1\rangle$ BEC) secures coherence during regeneration of the probe light pulse.

Revived light pulses under various conditions are shown in Fig. 2a–d. Whereas the BECs in Fig. 2a and b are condensed in separate potential wells, those in Fig. 2c and d are formed by adiabatically separating an already-formed condensate. We observe no qualitative difference between revivals in a single condensate (Fig. 3a), in adiabatically formed condensate-pairs (Fig. 2c, d), or in condensates that have always been separate (Fig. 2a). In all cases, no atoms are observed when we selectively image state $|2\rangle$ after the probe light pulse has been revived, indicating that the messenger atom pulse has been fully converted to light and state $|1\rangle$ atoms. No revived light pulse is observed when an isolated messenger atom pulse is illuminated with the coupling field (Fig. 2e).

It should be stressed that for light pulse revival to succeed with two distinct atom clouds as described, the atoms in both clouds must be Bose-condensed. The rate of coherent emission events in the revival process is determined by the Bose stimulation factor, $\hat{\psi}_1^\dagger \hat{\psi}_1$, for scattering into the second condensate. For a macroscopic occupation of this condensate, the stimulated processes completely dominate the spontaneous ones. By contrast, if the atoms formed a degenerate Fermi gas, attempts at revival in the second cloud would lead to emission rates below even the spontaneous rate obtained from a non-condensed, bosonic cloud[27,28]. It should also be noted that the two independently created condensates used in the experiments for light pulse storage and revival have a completely random relative phase. Therefore, interference experiments in which the revived light pulse interferes with a reference pulse would lead to high-contrast interference in each shot, but with random absolute fringe position.

Figure 2b shows a bimodal revival pulse obtained under the same conditions as in Fig. 2a, except that we let the messenger atom pulse propagate to a different location in the second $|1\rangle$ condensate before the light pulse is revived. Between storage and revival times, coherent atom dynamics create phase gradient differences and a different

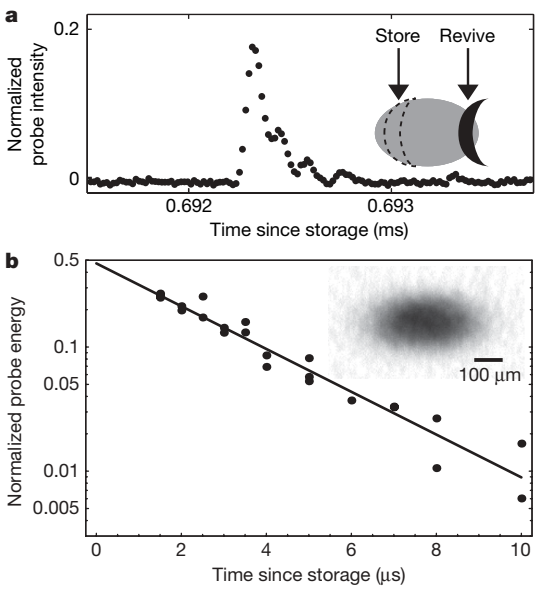


Figure 3 | Light pulse revivals in single clouds. **a**, A light pulse is revived at the right end of a single BEC of 3.4×10^6 atoms, 0.69 ms after storage in the left end. The cloud was prepared in a harmonic magnetic trap ($\omega_z = 2\pi \times 20$ Hz, $\omega_r = 2\omega_z$). All light parameters are similar to those in Fig. 2. **b**, Decay of revival signal in thermal cloud. The energy of the revived light pulse is plotted as a function of time since pulse storage. The thermal cloud (inset) has 13.5×10^6 atoms at 470 nK, just above the critical temperature for BEC (340 nK) in the trap described in **a**. Input probe Rabi frequency is $2\pi \times 3.2$ MHz; coupling storage and revival Rabi frequencies are $2\pi \times 3.5$ and $2\pi \times 17.5$ MHz, respectively.

position-dependent density ratio between the messenger and second $|1\rangle$ BEC. This determines the structure of the revived light pulse, as we have confirmed by numerical simulations.

These observations demonstrate coherent processing of optical information. In the experiments, expansion dynamics due to repulsive atom–atom interactions in the condensates after trap turn-off create $\sim 0.5\,\mathrm{rad}\,\mu\mathrm{m}^{-1}$ phase variations during the storage time. As controlling the revival time to within tens of microseconds controls the propagation depth of the messenger atom pulse to micrometre precision, we can revive the light pulse at locations where the phase patterns of the messenger and second BEC match. This leads to revived light pulses with the same shape as the incoming light pulse (Fig. 2a). For other propagation distances of the messenger pulse, various phase patterns can be imprinted on the revived light pulse, and differently shaped revival pulses are recorded at the photomultiplier. This coherent processing could be delicately controlled in trapped condensates, for example, by way of manipulation of atomic scattering lengths with Feshbach resonances[17].

Loss of $|2\rangle$ amplitude in the messenger pulse from atom–atom scattering[29,30], together with a roughly 50% loss from light pulse propagation in the condensate before storage, account for the difference in energy between the incident and the revived light pulses; there are no detectable losses from the storage and revival processes themselves. By careful selection of atomic species, magnetic sublevels, and manipulation of scattering lengths, we could minimize both slow light and atom pulse propagation losses. Shaping the density profile of the $|1\rangle$ condensates could increase the optical bandwidth of the process and further minimize losses.

We have observed the retrieval of optical information from a BEC after optical storage in a completely separate BEC. This is a result of slow-light-mediated atomic matter-wave amplification, which fully converts and coherently adds a messenger pulse of state $|2\rangle$ atomic amplitude to a receiving $|1\rangle$ condensate. Coherent atom pulses could repeatedly be moved from one condensate to another, which could be used as a replenishing scheme in a continuous-wave atom laser. The system also forms the basis for a new type of interferometer where spatially selected parts of an atomic wavefunction interfere. As demonstrated for bosons, such interferometry can be used for measurements of off-diagonal long-range order in spatially selected regions of degenerate gases of fermions and bosons.

We have demonstrated coherent optical information processing with matter wave dynamics: optical information is imprinted on a Bose-condensed atom cloud, quantum coherent atom dynamics alter the atomic imprint during the storage time, and finally the result is written back onto propagating optical fields. This points to a number of avenues in classical and quantum information processing. As a messenger atom pulse that embodies the incident light travels in free space, it can be independently trapped—potentially for minutes—and manipulated with external fields. Resulting classical and quantum states of the atom pulse can then be mapped onto revived light fields[17]. With the large optical delay-bandwidth products of $\sim 10^3$–10^4 already obtained here, this could also lead to novel designs for dynamically controllable optical delay lines.

Received 23 October; accepted 24 November 2006.

1. Southwell, K. (ed.) Ultracold matter. *Nature* **416,** 205–246 (2002).
2. Svistunov, B. V. & Shlyapnikov, G. V. Effect of Bose condensation on resonant optics in gaseous $H_\downarrow$. *Sov. Phys. JETP* **71,** 71–76 (1990).
3. Politzer, H. D. Light incident on a Bose-condensed gas. *Phys. Rev. A* **43,** 6444–6446 (1991).
4. Javanainen, J. Optical signatures of a tightly confined Bose condensate. *Phys. Rev. Lett.* **72,** 2375–2378 (1994).
5. You, L., Lewenstein, M. & Cooper, J. Line shapes for light scattered from Bose-Einstein condensates. *Phys. Rev. A* **50,** R3565–R3568 (1994).
6. Morice, O., Castin, Y. & Dalibard, J. Refractive index of a dilute Bose gas. *Phys. Rev. A* **51,** 3896–3901 (1995).
7. Hau, L. V., Harris, S. E., Dutton, Z. & Behroozi, C. H. Light speed reduction to 17 metres per second in an ultracold atomic gas. *Nature* **397,** 594–598 (1999).
8. Cornell, E. A. & Wieman, C. E. Nobel lecture: Bose-Einstein condensation in a dilute gas, the first 70 years and some recent experiments. *Rev. Mod. Phys.* **74,** 875–893 (2002).
9. Ketterle, W. Nobel lecture: When atoms behave as waves: Bose-Einstein condensation and the atom laser. *Rev. Mod. Phys.* **74,** 1131–1151 (2002).
10. Leggett, A. J. Bose-Einstein condensation in the alkali gases: Some fundamental concepts. *Rev. Mod. Phys.* **73,** 307–356 (2001).
11. Arimondo, E. Coherent population trapping in laser spectroscopy. *Prog. Opt.* **35,** 257–354 (1996).
12. Harris, S. E. Electromagnetically induced transparency. *Phys. Today* **50,** 36–42 (1997).
13. Knight, P. L., Stoicheff, B. & Walls, D. (eds) Highlights in quantum optics. *Phil. Trans. R. Soc. Lond. A* **355,** 2215–2416 (1997).
14. Kash, M. M. *et al.* Ultraslow group velocity and enhanced nonlinear optical effects in a coherently driven hot atomic gas. *Phys. Rev. Lett.* **82,** 5229–5232 (1999).
15. Budker, D., Kimball, D. F., Rochester, S. M. & Yashchuk, V. V. Nonlinear magneto-optics and reduced group velocity of light in atomic vapor with slow ground state relaxation. *Phys. Rev. Lett.* **83,** 1767–1770 (1999).
16. Liu, C., Dutton, Z., Behroozi, C. H. & Hau, L. V. Observation of coherent optical information storage in an atomic medium using halted light pulses. *Nature* **409,** 490–493 (2001).
17. Dutton, Z. & Hau, L. V. Storing and processing optical information with ultraslow light in Bose-Einstein condensates. *Phys. Rev. A* **70,** 053831 (2004).
18. Phillips, D. F., Fleischhauer, A., Mair, A., Walsworth, R. L. & Lukin, M. D. Storage of light in atomic vapor. *Phys. Rev. Lett.* **86,** 783–786 (2001).
19. Dutton, Z., Ginsberg, N. S., Slowe, C. & Hau, L. V. The art of taming light: Ultra-slow and stopped light. *Europhys. News* **35,** 33–38 (2004).
20. Baym, G. *Lectures on Quantum Mechanics* 425–426 (Benjamin/Cummings Publishing Co., Reading, Massachusetts, 1974).
21. Pitaevskii, L. & Stringari, S. *Bose-Einstein Condensation* (Oxford Univ. Press, Oxford, UK, 2003).
22. Lewenstein, M., You, L., Cooper, J. & Burnett, K. Quantum field theory of atoms interacting with photons: Foundations. *Phys. Rev. A* **50,** 2207–2231 (1994).
23. Cohen-Tannoudji, C., Dupont-Roc, J. & Grynberg, G. *Photons & Atoms* (Wiley, New York, 1997).
24. Deng, L. *et al.* Four-wave mixing with matter waves. *Nature* **398,** 218–220 (1999).
25. Inouye, S. *et al.* Phase-coherent amplification of atomic matter waves. *Nature* **402,** 641–644 (1999).
26. Kozuma, M. *et al.* Phase-coherent amplification of matter waves. *Science* **286,** 2309–2312 (1999).
27. Moore, M. G. & Meystre, P. Atomic four-wave mixing: Fermions versus bosons. *Phys. Rev. Lett.* **86,** 4199–4202 (2001).
28. Ketterle, W. & Inouye, S. Does matter wave amplification work for fermions? *Phys. Rev. Lett.* **86,** 4203–4206 (2001).
29. Burke, J. P., Greene, C. H. & Bohn, J. L. Multichannel cold collisions: Simple dependences on energy and magnetic field. *Phys. Rev. Lett.* **81,** 3355–3358 (1998).
30. Band, Y. B., Trippenbach, M., Burke, J. P. & Julienne, P. S. Elastic scattering loss of atoms from colliding Bose-Einstein condensate wave packets. *Phys. Rev. Lett.* **84,** 5462–5465 (2000).

Acknowledgements We thank J. Golovchenko and M. Burns for discussions, and W. Hill, Z. Dutton and J. MacArthur for technical assistance. This work was supported by the Air Force Office of Sponsored Research, the National Science Foundation, and the National Aeronautics and Space Administration.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to L.V.H. (hau@physics.harvard.edu).

LETTERS

Molecular fingerprinting with the resolved modes of a femtosecond laser frequency comb

Scott A. Diddams[1], Leo Hollberg[1] & Vela Mbele[1,2,3]

The control of the broadband frequency comb[1] emitted from a mode-locked femtosecond laser has permitted a wide range of scientific and technological advances—ranging from the counting of optical cycles for next-generation atomic clocks[1,2] to measurements of phase-sensitive high-field processes[3]. A unique advantage of the stabilized frequency comb is that it provides, in a single laser beam, about a million optical modes with very narrow linewidths[4] and absolute frequency positions known to better than one part in 10^{15} (ref. 5). One important application of this vast array of highly coherent optical fields is precision spectroscopy, in which a large number of modes can be used to map internal atomic energy structure and dynamics[6,7]. However, an efficient means of simultaneously identifying, addressing and measuring the amplitude or relative phase of individual modes has not existed. Here we use a high-resolution disperser[8,9] to separate the individual modes of a stabilized frequency comb into a two-dimensional array in the image plane of the spectrometer. We illustrate the power of this technique for high-resolution spectral fingerprinting of molecular iodine vapour, acquiring in a few milliseconds absorption images covering over 6 THz of bandwidth with high frequency resolution. Our technique for direct and parallel accessing of stabilized frequency comb modes could find application in high-bandwidth spread-spectrum communications with increased security, high-resolution coherent quantum control, and arbitrary optical waveform synthesis[10] with control at the optical radian level.

The proposition of using the frequency comb from a mode-locked laser for optical spectroscopy has existed for at least three decades. The proposal and early experiments of groups in Novosibirsk and Stanford[11,12] highlighted the advantages of a mode-locked laser frequency comb for two-photon spectroscopy. The more recent developments in stabilization of the carrier-envelope offset frequency[1,13] provide powerful new tools that have led to recent advances in high-resolution ($\sim 10^{-11}$) spectroscopy performed directly with the output of the mode-locked laser[6,7]. In all such experiments, the various atomic systems under study acted as the high-resolution spectral discriminator that effectively selected an individual frequency comb element (or groups of comb elements) out of a greater number of elements that passed through the sample. Here we take a significantly different approach, in which a high-resolution spectrometer is used to spatially separate and resolve individual comb elements in a

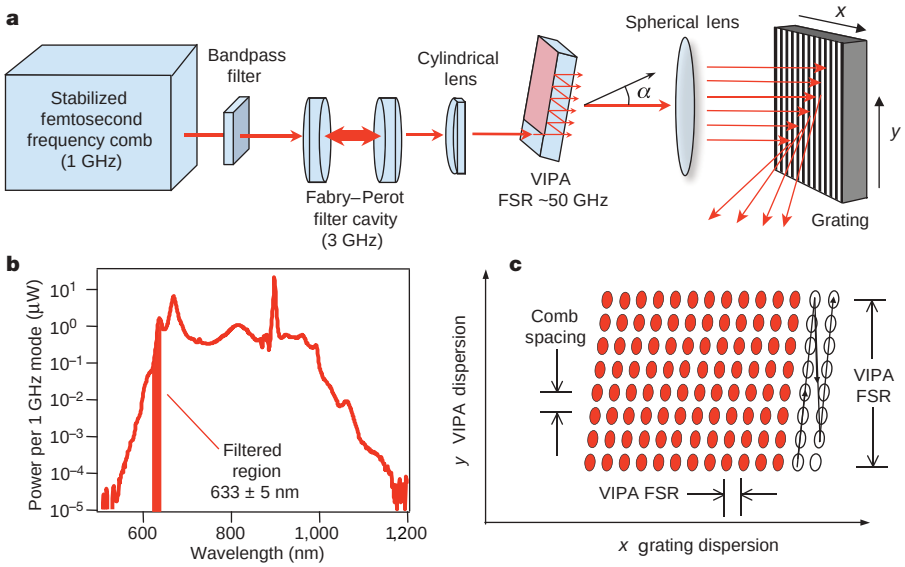


Figure 1 | Experimental set-up. a, A high-resolution virtually imaged phased array (VIPA) disperser is used in combination with a diffraction grating to spatially resolve the stabilized frequency comb of a Ti:sapphire femtosecond laser. The full output spectrum of the laser and the 633 nm region isolated by the bandpass filter are shown in **b**. The spectrometer output consists of a two-dimensional array of the frequency comb modes, where each 'dot' represents an individual mode (**c**). Within a column (y), which is tilted by the grating dispersion, the dots are separated by the mode spacing (3 GHz in this case). Within each row (x), the dots are separated by the VIPA free spectral range (FSR, ~50 GHz in this case)[17]. The manner in which successive modes can be indexed and counted is indicated by the arrows in the rightmost two columns. For clarity, not all modes are shown in this diagram.

[1]Time and Frequency Division, National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA. [2]CSIR, NML, PO Box 395, Pretoria, ZA-0001, South Africa. [3]School of Physics, University of the Witwatersrand, Johannesburg, ZA-2050, South Africa.

parallel architecture. Thus, a multi-channel detector can measure the amplitude of the individual comb elements. Using a fraction of the input field as a reference, the relative phase shift introduced by a resonant transition could also be measured with an approach similar to spectral interferometry[14].

Our approach is shown schematically in Fig. 1a. The frequency comb is produced with a broadband Ti:sapphire laser having a repetition rate $f_{rep} = 1$ GHz (refs 15, 16). Both the carrier-envelope offset frequency (f_0) and f_{rep} are stabilized to a low-noise microwave frequency standard (a hydrogen maser), such that the frequency of each element of the comb may be determined absolutely with a fractional uncertainty at or below $\sim 2 \times 10^{-13}$ for averaging times of 1 s and longer. Although an optical reference would provide ~1 Hz optical resolution[4] and fractional uncertainty into the 10^{-17} range[5], the microwave reference is sufficient for these experiments. Indeed, for most spectroscopy experiments we could imagine, the atomic reference provided freely by the global positioning system (GPS) would be more than adequate. The output of the Ti:sapphire laser spans roughly 600–1,000 nm, providing tremendous bandwidth for spectroscopic measurements (Fig. 1b). However, imaging and recording the full bandwidth in a single measurement proved challenging. Thus we use an optical bandpass filter to restrict the spectrum to 10 nm of bandwidth around 633 nm, which still provides ~75 billion resolvable spectral bins in 1 s of averaging. Within this 10 nm bandwidth, a Fabry–Perot cavity (finesse ~300) consisting of two spherical mirrors is used to further filter (thin) the frequency comb to a mode spacing of $3f_{rep}$, which better matches the resolution of the spectrometer that follows. By changing the length of the Fabry–Perot cavity, we have filtered the comb at integer multiples of the repetition rate up to $14f_{rep}$. Such flexible filtering of the comb is useful not only in spectroscopy, but also for waveform synthesis and communications applications.

The high-resolution spectrometer (Fig. 1a) provides ~1.2 GHz resolution in the visible (633 nm) spectral region by combining a virtually imaged phased array (VIPA) spectral disperser[8] with a conventional grating in an orthogonal arrangement[9,17]. The VIPA is essentially a plane-parallel solid etalon, where the input beam (focused to a line) is injected at an angle through an uncoated entrance window on the front face. The remainder of the front face is coated with a high-reflective dielectric coating, while the back face has a dielectric coating with 96% reflectivity. The multiple reflections within the VIPA etalon interfere such that the exiting beam has its different frequencies emerging at different angles. As with all etalons, the VIPA has a free spectral range (FSR; ~50 GHz in this case) determined by its thickness and material index of refraction. The result is that for an input with spectral bandwidth greater than 50 GHz, the output orders are spatially superimposed on each other. This problem is well known in classical spectroscopy, and has been overcome by using a second dispersive element along an orthogonal spatial dimension[18]. Recent implementations, directed towards separating densely spaced optical communications channels, use a diffraction grating orthogonal to the VIPA[9,17]. In such a case, the grating should provide spectral resolution better than that of the VIPA's FSR. We achieve ~20 GHz resolution for visible light with a 2,400 lines mm^{-1} grating used at a large angle of incidence, such that ~24,000 lines of the grating are illuminated.

The output of the VIPA/grating spectrometer is imaged onto a charge-coupled device (CCD) camera (6.7 μm pixel pitch), resulting in an array of 'dots', representing the power of the individual comb modes spaced by 3 GHz. Thus, the spectrometer transfers the one-dimensional comb into something more reminiscent of a two-dimensional 'brush'. This is illustrated in Fig. 1c, while actual data are shown in Fig. 2a. In the vertical direction of this image, the data repeats every 50 GHz at the FSR of the VIPA, and a subset of unique data are enclosed by the black boxes superimposed on the data of Fig. 2. Inside this boundary, almost 2,200 individual modes can be clearly resolved, spanning the ~6.5 THz bandwidth captured on the CCD. The skew of the columns of modes results from the particular choice of the 3 GHz mode spacing, the FSR of the VIPA, the angular dispersion of the grating, and additionally, a slight rotation of the experimental apparatus relative to the camera axes. The repetitive nature of the data in the vertical direction is more evident in Fig. 2b, which is an image acquired in 5 ms with an iodine vapour cell inserted in the beam path before the spectrometer. The cell is at room temperature (25 °C), and multi-passed to yield an equivalent length of ~2 m. As seen, numerous modes are attenuated owing to their

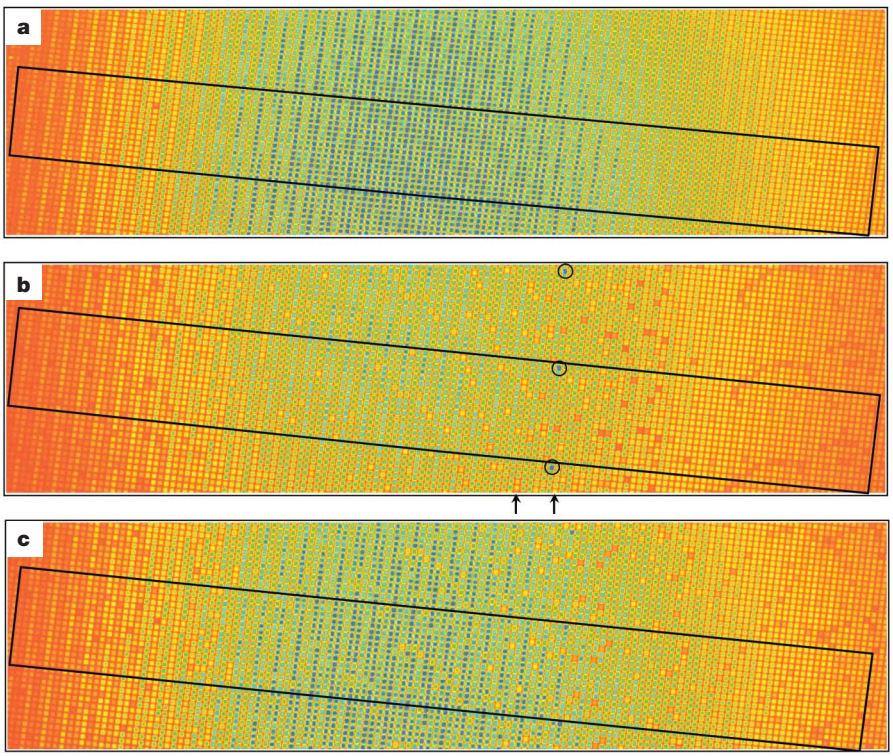


Figure 2 | Two-dimensional spectrograms of optical frequency 'brush'. The array of 'dots' is the false-colour image of ~2,200 individual modes spanning 6.5 THz. (For the chosen colour map, blue represents higher optical power; a standard dilation filter was applied to aid visual clarity, and this causes the individual modes to have a square-like appearance.) In the vertical direction of this image, the data repeat every 50 GHz, and a subset of unique data are enclosed by the black box superimposed on the data. **a**, Background image taken without iodine present. **b**, Image acquired with iodine present. Numerous modes are attenuated owing to absorption by the iodine vapour, thus providing a unique fingerprint of the molecular vapour. The circled modes are those of a reference laser used in the calibration of the data. **c**, Same as **b** but with a different value of f_{rep} for the frequency comb, revealing a different set of absorbing transitions.

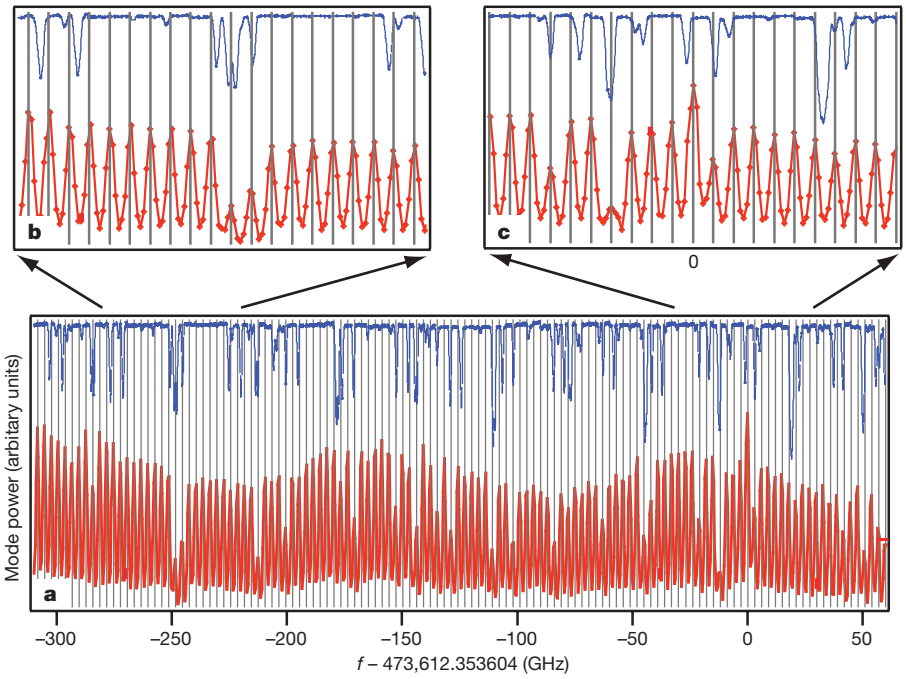


Figure 3 | Concatenated line spectra. **a**, Adjacent line spectra from seven columns identified by arrows in Fig. 2b are concatenated to show the power detected from >100 individual modes (red traces, bottom axis). No dilation filtering was applied in this case. Modes that are attenuated owing to the absorption of the iodine vapour appear with smaller amplitude relative to adjacent modes. The slow (~150 GHz) amplitude modulation arises from imperfections in the imaging and is of no physical consequence. Vertical grey lines identify the calculated mode frequencies, while the blue trace along the upper axis is the measured transmission spectra of iodine using CW laser techniques[20]. **b**, **c**, Expanded sections from **a** showing the attenuation of specific modes, the 1.2 GHz resolution limit of the spectrometer, and the pixelization of the CCD camera (small points on the red trace).

coincidence with various absorbing transitions in the iodine. At room temperature, the Doppler-broadened linewidth of the iodine transitions is ~400 MHz; however, the hyperfine splitting causes additional broadening of the lines to ~1–2 GHz.

An advantage of using the self-referenced frequency-stabilized comb for this spectroscopy is that the frequencies of the modes emitted from the laser are absolutely fixed according to $\nu_n = nf_{rep} + f_0$, where n is the integer mode index ($n \approx 500{,}000$). In principle, each pixel of the CCD image then can be assigned a unique frequency. However, this requires that n be determined for each of the imaged modes, keeping in mind that only one-third of the original modes are transmitted by the filter cavity. In this example, the calibration is accomplished with the aid of a continuous wave (CW) He–Ne laser stabilized to the well-known a_{16} component (line 'f') of the R(127)11–5 transition in $^{127}I_2$, with frequency equal to 474,612.353604 GHz (ref. 19). This laser is overlapped with the frequency comb in a single-mode optical fibre, and sent through the same high-resolution spectrometer. Its output is visible in Fig. 2b as the three noticeably more intense modes (corresponding to three orders of the VIPA) that are circled. Simultaneous heterodyne measurement between the He–Ne laser and the filtered frequency comb in a separate high-speed detector yields unique identification of the mode index n and absolute frequency calibration. Following the labelling technique shown diagrammatically in Fig. 1c, adjacent columns consisting of 16 modes can then be concatenated to display the data on a more traditional linear frequency axis. This is done for the

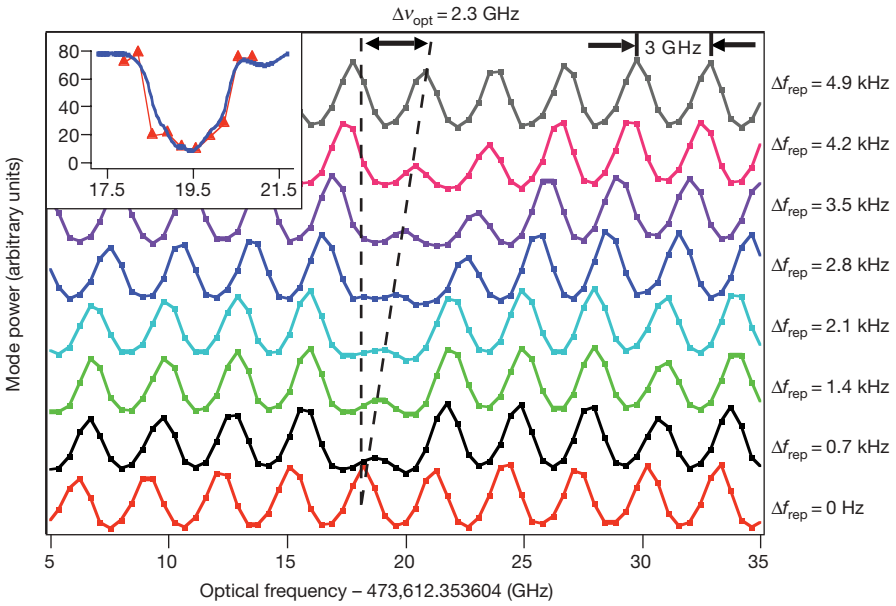


Figure 4 | Absorption spectra of P(32)6–3, R(59)8–4 and R(53)8–4 transitions in iodine. Eight line spectra, each obtained from images similar to Fig. 2, are offset vertically for clarity. In the successive spectra, the repetition rate (f_{rep}) was increased by 700 Hz. As a result, a specific optical mode is scanned a total of $\Delta\nu_{opt} = 2.3$ GHz with ~330 MHz steps across the absorption feature due to the overlapping P(32)6–3, R(59)8–4 and R(53)8–4 transitions[22]. The inset shows the de-convolved absorption feature (red triangles representing the relative transmission through the iodine versus the same optical frequency of the abscissa of the main plot), obtained from a fit to the peak value of the attenuated mode. The result of CW laser spectroscopy (blue line)[20] is shown for comparison.

seven columns between the arrows at the bottom of Fig. 2b, and the result is given in Fig. 3. The lower trace (red) in each of the plots of Fig. 3 is obtained from concatenating line traces along the unprocessed CCD image. Along the top axis, we also plot (blue) the iodine transmission spectra obtained from CW laser spectroscopy[20]. The calculated positions of the comb modes are indicated by the vertical (grey) lines. Only modes that coincide with an iodine transition are attenuated, thus demonstrating the potential for this technique to yield quantitative and accurate results.

The width of the intensity peaks in Fig. 3 data show the spectrometer resolution to be ~1.2 GHz at 474 THz, which is within a factor of two of the calculated resolution. This is well-matched to the pixel resolution of the camera, which corresponds to about 440 MHz. However, this should not be confused with the linewidth of the comb modes, which is ~100 kHz with the present maser reference, but could be significantly smaller (~1 Hz) with an optical reference for the frequency comb. Within the 1.2 GHz spectrometer resolution, we estimate that the absolute frequency position of an individual comb element is determined to 20 kHz in 5 ms of averaging, decreasing to 100 Hz at 1 s. Scanning the repetition rate of the laser (with the filter cavity also tracking) enables one to scan out the full optical spectrum with a resolution suitable for the system under study. In Fig. 2c a second image of the same spectral region but with a different repetition rate is shown. Visual differences with Fig. 2b are clearly seen. More quantitative results are obtained by again comparing line traces from images acquired with different repetition rates. This is shown in Fig. 4, where line traces covering 30 GHz are displayed for eight different values of f_{rep} offset sequentially by 700 Hz—corresponding to a 330 MHz offset from trace-to-trace in the optical domain. In this manner, a specific absorption feature can be mapped. Rapid electronic tuning of the repetition rate under computer control would yield a complete scan over many THz with appropriate resolution (~200 MHz) for Doppler-limited spectroscopy at video rates.

Although the straightforward absorption spectroscopy performed here is sufficient for strong transitions, the high resolution of this approach would benefit from cavity-enhanced sensitivity (achieved by, for example, putting the vapour cell inside an optical cavity), as has been demonstrated in the case of broadband cavity ringdown spectroscopy[21]. Additionally, multiple spectrometers at various wavelengths could be combined with image correlation techniques for identification of species of interest. In fact, the ability to spatially isolate and detect the individual modes of the stabilized frequency comb will give a way to implement (in a massively parallel manner) almost any of the powerful CW laser spectroscopic techniques that have been developed.

Received 23 September; accepted 12 December 2006.

1. Udem, T., Holzwarth, R. & Hansch, T. W. Optical frequency metrology. *Nature* **416,** 233–237 (2002).
2. Diddams, S. A. *et al.* An optical clock based on a single trapped $^{199}Hg^{+}$ ion. *Science* **293,** 825–828 (2001).
3. Baltuska, A. *et al.* Attosecond control of electronic processes by intense light fields. *Nature* **421,** 611–615 (2003).
4. Bartels, A., Oates, C. W., Hollberg, L. & Diddams, S. A. Stabilization of femtosecond laser frequency combs with subhertz residual linewidths. *Opt. Lett.* **29,** 1081–1083 (2004).
5. Oskay, W. H. *et al.* Single-atom optical clock with high accuracy. *Phys. Rev. Lett.* **97,** 020801 (2006).
6. Marian, A., Stowe, M. C., Lawall, J. R., Felinto, D. & Ye, J. United time-frequency spectroscopy for dynamics and global structure. *Science* **306,** 2063–2068 (2004).
7. Gerginov, V., Tanner, C. E., Diddams, S. A., Bartels, A. & Hollberg, L. High-resolution spectroscopy with a femtosecond laser frequency comb. *Opt. Lett.* **30,** 1734–1736 (2005).
8. Shirasaki, M. Large angular dispersion by a virtually imaged phased array and its application to a wavelength demultiplexer. *Opt. Lett.* **21,** 366–368 (1996).
9. Xiao, S. & Weiner, A. M. 2-D wavelength demultiplexer with potential for ≥1000 channels in the C-band. *Opt. Express* **12,** 2895–2901 (2004).
10. Jiang, Z., Seo, D. S., Leaird, D. E. & Weiner, A. M. Spectral line-by-line pulse shaping. *Opt. Lett.* **30,** 1557–1559 (2005).
11. Baklanov, Y. V. & Chebotayev, V. P. Narrow resonances of two-photon absorption of super-narrow pulses in a gas. *Appl. Phys.* **12,** 97–99 (1977).
12. Teets, R., Eckstein, J. & Hänsch, T. W. Coherent two-photon excitation by multiple light pulses. *Phys. Rev. Lett.* **38,** 760–764 (1977).
13. Jones, D. J. *et al.* Carrier-envelope phase control of femtosecond mode-locked lasers and direct optical frequency synthesis. *Science* **288,** 635–639 (2000).
14. Lepetit, L., Cheriaux, G. & Joffre, M. Linear techniques of phase measurement by femtosecond spectral interferometry for applications in spectroscopy. *J. Opt. Soc. Am. B* **12,** 2467–2474 (1995).
15. Bartels, A. & Kurz, H. Generation of a broadband continuum by a Ti:sapphire femtosecond oscillator with a 1-GHz repetition rate. *Opt. Lett.* **27,** 1839–1841 (2002).
16. Ramond, T. M., Diddams, S. A., Hollberg, L. & Bartels, A. Phase-coherent link from optical to microwave frequencies by means of the broadband continuum from a 1-GHz Ti:sapphire femtosecond oscillator. *Opt. Lett.* **27,** 1842–1844 (2002).
17. Wang, S. X., Xiao, S. & Weiner, A. M. Broadband, high spectral resolution 2-D wavelength-parallel polarimeter for dense WDM systems. *Opt. Express* **13,** 9374–9380 (2005).
18. Jenkins, F. A. & White, H. E. *Fundamentals of Optics* 4th edn (McGraw Hill, New York, 1976).
19. Quinn, T. J. Practical realization of the definition of the metre, including recommended radiations of other optical frequency standards (2001). *Metrologia* **40,** 103–133 (2003).
20. Kato, H. *Doppler-Free High Resolution Spectral Atlas of Iodine Molecule 15000 to 19000 cm^{-1}* (Japan Society for the Promotion of Science, Tokyo, 2000).
21. Thorpe, M. J., Moll, K. D., Jones, R. J., Safdi, B. & Ye, J. Broadband cavity ringdown spectroscopy for sensitive and rapid molecular detection. *Science* **311,** 1595–1599 (2006).
22. Simonsen, H. R. & Rose, F. Absolute measurements of the hyperfine splittings of six molecular $^{127}I_2$ around the He-Ne/I_2 wavelength at λ≈633 nm. *Metrologia* **37,** 651–658 (2000).

Acknowledgements We thank A. M. Weiner for discussions about the VIPA spectrometer, and J. Ye for discussions that motivated our application of frequency combs to broadband spectroscopy. We further thank J. Stalnaker and Y. LeCoq for their comments on this manuscript, and H. Kato for the CW iodine spectroscopy data. This paper is a contribution of the National Institute of Standards and Technology, with partial support from DARPA.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.A.D. (sdiddams@boulder.nist.gov).

LETTERS

Patterning of sodium ions and the control of electrons in sodium cobaltate

M. Roger[1], D. J. P. Morris[2], D. A. Tennant[3,4], M. J. Gutmann[5], J. P. Goff[2], J.-U. Hoffmann[3], R. Feyerherm[3], E. Dudzik[3], D. Prabhakaran[6], A. T. Boothroyd[6], N. Shannon[7], B. Lake[3,4] & P. P. Deen[8]

Sodium cobaltate (Na_xCoO_2) has emerged as a material of exceptional scientific interest due to the potential for thermoelectric applications[1,2], and because the strong interplay between the magnetic and superconducting properties has led to close comparisons with the physics of the superconducting copper oxides[3]. The density x of the sodium in the intercalation layers can be altered electrochemically, directly changing the number of conduction electrons on the triangular Co layers[4]. Recent electron diffraction measurements reveal a kaleidoscope of Na^+ ion patterns as a function of concentration[5]. Here we use single-crystal neutron diffraction supported by numerical simulations to determine the long-range three-dimensional superstructures of these ions. We show that the sodium ordering and its associated distortion field are governed by pure electrostatics, and that the organizational principle is the stabilization of charge droplets that order long range at some simple fractional fillings. Our results provide a good starting point to understand the electronic properties in terms of a Hubbard hamiltonian[6] that takes into account the electrostatic potential from the Na superstructures. The resulting depth of potential wells in the Co layer is greater than the single-particle hopping kinetic energy and as a consequence, holes preferentially occupy the lowest potential regions. Thus we conclude that the Na^+ ion patterning has a decisive role in the transport and magnetic properties.

Within the metallic CoO_2 units, the state of the conduction electrons is unconventional, with large effective mass and small Fermi temperature, a high Seebeck coefficient[1], and superconductivity below 5 K when hydrated[3]. Furthermore, magnetic fields strongly influence the thermopower and it has been proposed that large spin entropy, associated with the conduction electrons, causes cooling when an electric current flows under applied voltage[7]. Although the exact origin of these exotic magnetic and transport properties remains contentious, both the spin and electronic degrees of freedom of Co ions and the layered hexagonal geometry of CoO_2 are of key importance. Another remarkable feature of Na_xCoO_2 is that Na^+ ions sandwiched between CoO_2 layers are able to tunnel between different intercalation sites. The Na^+ layers are chargeable electrochemically and the plateaus and inflections, measured in capacity-voltage[8], indicate complex behaviour. At present it is not clear whether the mechanism for sodium ordering relies mainly on Coulomb interactions[9], or on more complex cooperative phenomena involving charge or orbital ordering in the Co layers[10].

Sodium ions can occupy sites between oxygen atoms in the CoO_2 layers, as illustrated in Fig. 1a. These form two interpenetrating triangular lattices, denoted Na1 and Na2. Cobalt ions lie above and below Na1 sites, resulting in an extra energy cost for occupation relative to Na2 of Δ, which depends on the effective dielectric constant ε. Na^+ ions are bigger than the distance between Na1 and Na2 sites, disallowing simultaneous occupancy of nearest-neighbouring sites. In addition there is electrostatic repulsion between the Na^+ ions consisting of a long-range Coulomb potential $e/(4\pi\varepsilon\varepsilon_0 r)$ and a short-range Na–Na ion shell repulsion $V = 0.06$eV between neighbours on the same sublattice.

Exact Ewald summations of long-range Coulomb energies reveal that the spontaneous formation of multi-vacancy clusters drives the organization of Na^+ ions, as illustrated in Fig. 1b. Na^+ ion vacancies become attractive at very short length scales, through a delicate balance of Coulomb forces and onsite Na1 energy, and condense into droplets. As droplet size increases, the extra energy cost of Na1 sites in the droplet core eventually cancels the energy gain of the droplet surface. Below a critical temperature and for simple fractional values of x, droplets of identical size condense in regular arrays, keeping them as far apart as possible to minimize Coulomb repulsion. Above this temperature, entropy is gained by the coexistence of different-sized droplets, keeping only short-range correlations. This 'glassy' droplet structure is destroyed at temperatures which exceed the stability domain of Na_xCoO_2.

At low concentration we find the stable phases at $x = 1/3$ comprising only Na2 sites, and $x = 1/2$ with an equal occupation of Na1 and Na2 sites, as observed experimentally[10,11]. This latter striped phase can also be viewed as closely packed di-vacancy clusters. Figure 1c compares the ground-state energy for increasing simple fractional concentrations x of regular arrays of ordered mono-, di-, tri- and quadri-vacancy clusters. Although di-vacancy phases are the most favourable for $0.5 < x < 0.71$, tri-vacancy phases gain more energy at $0.75 < x < 0.8$ and quadri-vacancy phases, although they are very close in energy to tri-vacancy phases, are marginally stable for $0.8 < x < 0.85$. There are a number of phases in the vicinity of $x \approx 0.8$ with comparable energy, and this explains the observation of a wide variety of superstructures in this composition range. In addition, for $x > 0.8$ coexistence of $x = 1$ phase with a phase of lower $x \approx 0.8$ is predicted, as observed through nuclear magnetic resonance (NMR)[12].

We performed neutron diffraction measurements on single-crystal Na_xCoO_2 of composition $x = 0.75$, 0.78 and 0.92; see Methods. Figure 2a–c shows the data for $x = 0.75$ at $T = 150$ K for the hexagonal $l = 11$, 10 and 9 planes, respectively, and this pattern is observed down to the lowest temperatures studied, $T = 1.5$ K. Rings of superlattice peaks form around the main Bragg peaks, together with higher-order harmonics, and these arise from

[1]Service de Physique de l'Etat Condensé, (CNRS/MIPPU/URA 2464), DSM/DRECAM/SPEC, CEA Saclay, P.C. 135, F-91191 Gif Sur Yvette, France. [2]Department of Physics, University of Liverpool, Oliver Lodge Laboratory, Liverpool L69 7ZE, UK. [3]Hahn-Meitner Institut, Glienicker Strasse 100, Berlin D-14109, Germany. [4]Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstrasse 36, Berlin D-10623, Germany. [5]ISIS Facility, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK. [6]Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK. [7]H. H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, UK. [8]European Synchrotron Radiation Facility, BP 220, F-38043 Grenoble cedex, France.

long-range ordering of Na ions. There is a complex modulation of superlattice peak intensity along l, and within the plane around the rings of superstructure peaks.

The observed commensurate modulation wavevectors correspond to the supercell of the $x = 0.80$ tri-vacancy model, $\mathbf{a}' = \mathbf{a} + 3\mathbf{b}$ and $\mathbf{b}' = 4\mathbf{a} - 3\mathbf{b}$ (see Supplementary Information Fig. 1), which we show to be within a few kelvin of the energetically most stable phase in Fig. 1c. The Fourier transforms from this phase show good agreement with experiment; see Fig. 2d–f. The variation of intensity within the plane rules out other mono- and di-vacancy cluster models; see Supplementary Information Fig. 2. The electrostatic potential energy is lowest when the tri-vacancy clusters in successive sodium layers are shifted by the maximum amount within the plane to the body centre of the unit cell, and the calculated scattering agrees best with the data for this model. To obtain quantitative agreement with the observed intensities, further displacement of CoO_6 octahedra is required. The gradient of the electrostatic potential from Na^+ ions has been calculated at each of the Co sites and there is a force component along the hexagonal c-direction away from the tri-vacancy clusters. The displacement of the Co ions is assumed to be proportional to this force, and the oxygen ions move in such a way to keep the Co–O bond length constant (Fig. 3). The model has just one adjustable parameter, the force constant, and best agreement with the data are obtained with a maximum displacement of the Co/O ions of only $0.01c$ ($\pm 0.001c$ error).

We observe a first-order phase transition to the predicted $x = 0.84$ quadri-vacancy phase of Fig. 1c above $T = 285$ K, and this coincides with an anomaly in the electrical resistance, see Supplementary Figs 3, 4 and 5. The superstructures have also been determined for $x = 0.78$ and 0.92, and these can be understood in terms of the coexistence of phases at low temperature and the quadri-vacancy phase at high temperature. The fractional occupation of Na sites calculated in

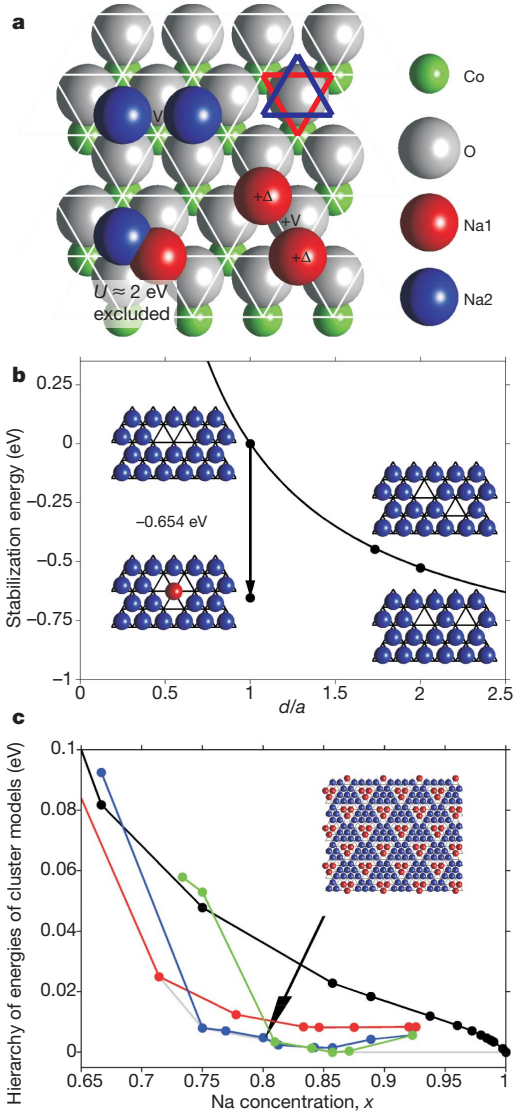


Figure 1 | Na_xCoO_2 has Na^+ ions intercalated between CoO_2 layers. **a**, Two interpenetrating hexagonal lattices of intercalation sites denoted Na1, and Na2 in a repeated star-of-David motif are formed. Ions are shown to scale. The energy 2 eV may be considered infinite at room temperature and, therefore, the occupation of neighbouring Na1 and Na2 sites is excluded, as described in Methods. **b**, Energy for two vacancies decreases with increasing distance d, in the units of the hexagonal lattice parameter a, as expected for Coulomb repulsion. Neighbouring vacancies can reduce their energy by promotion of a Na2 sodium to the central Na1 site. The resultant di-vacancy cluster has net charge $2e^-$ spread over three sites and substantially lower energy, because the central sodium is now further from its neighbours. Its stabilization energy is lower than vacancies three to four sites apart, and so should form spontaneously at modest concentrations. The formation of tri-vacancies and quadri-vacancies follows a similar process. **c**, Ground-state energies of superstructures for mono-vacancies (black), di-vacancies (red), tri-vacancies (blue) and quadri-vacancies (green); a function is subtracted to show energy differences on a milli-electronvolt scale. The inset shows the $x = 0.80$ tri-vacancy phase.

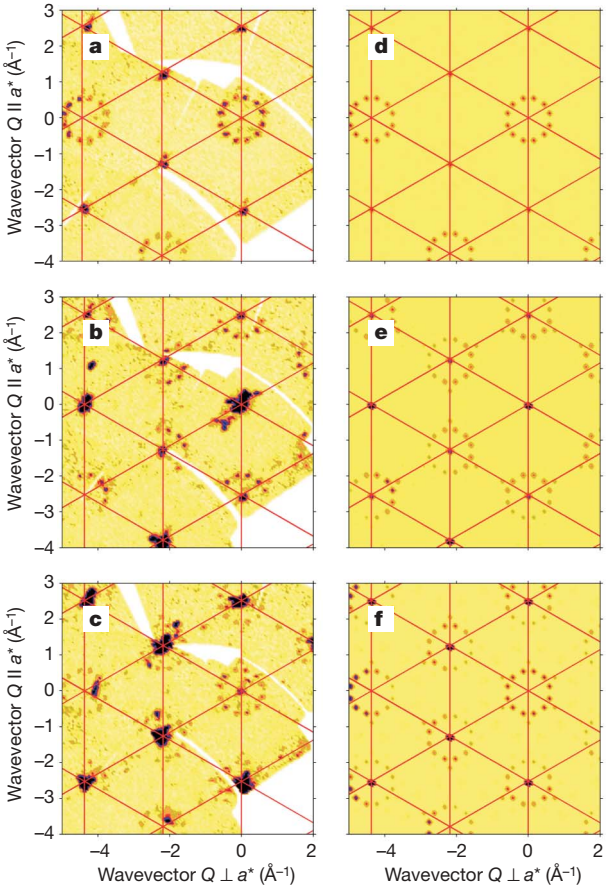


Figure 2 | Neutron diffraction showing bulk three-dimensional ordering of Na^+. Q is wavevector transfer. The planes shown are: **a**, $(h,k,11)$; **b**, $(h,k,10)$; **c**, $(h,k,9)$, for $x = 0.75$ at $T = 150$ K. Superlattice reflections form rings around hexagonal Bragg reflections, and some higher-order harmonics are observed. **d–f**, The corresponding scattering calculated using the $x = 0.80$ tri-vacancy phase shown in Fig. 3. The lattice is averaged over all domains. The scattering calculated using a one-parameter model (see text) captures the observed modulation both within and between planes. The cancellation of intensity due to interference between regular oxygen sites makes the scattering at $L = 11$ particularly sensitive to the sodium ordering and associated distortion field.

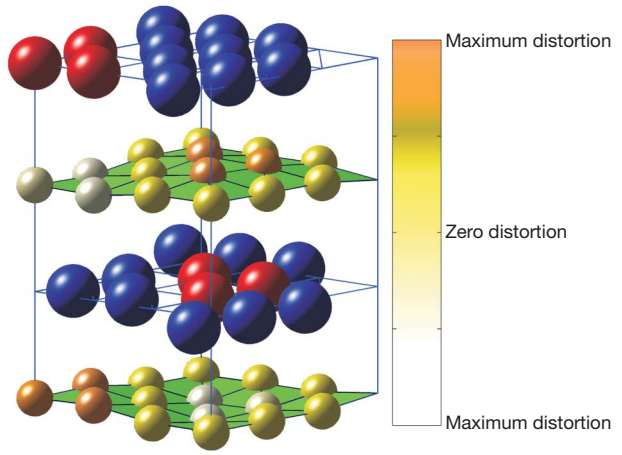


Figure 3 | Real space superstructure. The unit cell for the $x = 0.80$ tri-vacancy superstructure used to model the low temperature phase for $x = 0.75$, showing the in-plane displacement of tri-vacancies in successive Na layers. The colour scale on the Co ions shows their displacements, proportional to the electrostatic potential gradient. The maximum displacement of cobalt ions, $0.01c$, and the separation between planes are increased in this figure to aid visualization. The O ions are omitted for clarity, but their positions follow the Co ion distortion because the Co–O bond length remains fixed. This leads to a small in-plane displacement of the O ions, causing a small in-plane movement of Na ions (out-of-plane displacement of the Na ions is ruled out by symmetry).

our model agrees with the available neutron diffraction data[13–15] over a wide range of x (see Supplementary Fig. 6).

We clearly identify two energy scales: the electrostatic energy driving the formation of vacancy clusters (~500 meV), and a much lower energy from the deformation of the CoO_2 layer (a few millielectronvolts). The good agreement of the experimental results with the vacancy clustering predicted by our first-order treatment opens up the study of elastic deformations and cooperative interactions with mobile carriers in the CoO_2 layers using perturbative methods. This approach may be applied to other transition-metal oxides with ions partially filling spaces in hexagonal planes.

The Na^+ ion patterning determines the Coulomb landscape in the Co planes, and this depends very sensitively on the Na superstructure. The electronic properties may be modelled in terms of a modified Hubbard model

$$H = -t \sum_{\langle i,j\rangle\sigma=\uparrow,\downarrow} c^{+}_{i\sigma} c_{j\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow} + \sum_{i} E_i \left(n_{i\uparrow} + n_{i\downarrow}\right)$$

where U is the penalty for double occupation of a Co site, $c^{+}_{i\sigma}$ is a creation operator for holes, and the E_i values are on-site energies corresponding to different positions of the Co ions with respect to the modulation of the Coulomb landscape in the supercell, induced by the ordered array of Na ions.

Figure 4 shows two contrasting cases, $x = 0.5$ and 0.80. Coulomb-potential stripes are obtained for $x = 0.5$ in agreement with the low- and high-spin charge-ordered stripes found using neutron diffraction[16]. Potential wells of depth ~100 meV are predicted for $x = 0.80$. Because the total bandwidth of $Na_{0.7}CoO_2$ is less than 100 meV and the single-particle hopping frequency is about $t \approx 10$ meV (ref. 17), the electrostatic potential may be expected to localize holes. This picture is consistent with the multiple valence states detected by NMR[18], and the conduction paths around the localized Co^{3+} regions readily account for the observed metallic behaviour. Furthermore, the fact that the separation between localized regions of spin within the plane is comparable to the spacing between planes explains the three-dimensional magnetism revealed by the magnetic excitations[19,20].

Notably, multi-vacancy clusters form cages in which Na ions can vibrate relatively freely. These would be expected to disrupt the propagation of phonon excitations, leading to a low thermal

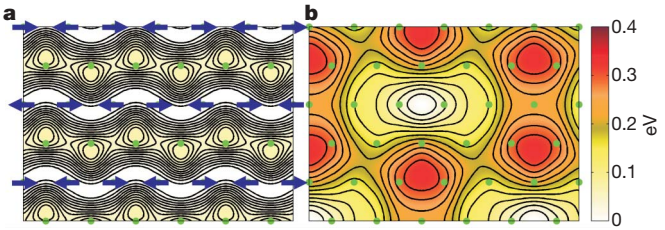


Figure 4 | Coulomb potential in the Co plane calculated using ordered superstructures. Positions of Co ions are shown as green circles, and the colour scale is in electronvolts. Sodium superstructures are: **a**, di-vacancy array for $x = 1/2$. The depth of the potential wells is of the order of 60 meV. Holes should localize along white stripes in Co planes. The observed striped magnetic order is shown as blue arrows[11]. **b**, Tri-vacancy array for $x = 0.80$, as shown in Fig. 3. Co^{4+} is expected at the minima and Co^{3+} at the maxima. Spin-half holes are localized in potential wells of depth 100 meV, where localized magnetic moments are expected.

conductivity. The 'rattling' of cations inside lattice cavities has been observed in a range of promising candidates for thermoelectric applications[21–23]. This behaviour is confirmed for metallic Na_xCoO_2 by the reported low, temperature-independent thermal conductivity[11]. The rare coincidence of high electrical conductivity and low thermal conductivity are precisely the conditions required for thermoelectric materials with high figures of merit.

Finally, while the mechanism of superconductivity in the hydrated phases will also depend upon other factors, sodium clustering will have an important effect. The electrostatic influence of Na ordering on the CoO_2 layers will be extremely sensitive to the intercalation of dipolar water molecules. In bi-hydrate $Na_xCoO_2 \cdot yH_2O$, water molecules intercalate between the Na^+ ions and the CoO_2 planes, and therefore screen the conduction electrons from the Na potential. Superconductivity is observed at 5 K (ref. 3), consistent with effective screening of the pair-breaking Coulomb potential. In contrast, for mono-hydrates the water molecules are intercalated within the planes of Na^+ ions and cannot screen the CoO_2 layers effectively, and indeed, like their anhydrous counterparts, these materials fail to superconduct[24].

In summary, elaborate sodium ordering in Na_xCoO_2 has been measured using neutron scattering. Organization of the sodium atoms proceeds by formation of vacancy clusters which order in a periodic lattice. We propose a model to explain the results, based on ionic shell and Coulomb interactions between sodium ions. The ordered sodium lattice influences the metallic properties of the CoO_2 layers via its periodic Coulomb potential and exerts control over many of the thermoelectric, magnetic and superconducting properties. Our results show that Na_xCoO_2 is a model material with which to investigate electrochemical control of magnetic and electronic properties.

METHODS

Experiment. Measurements were made from samples obtained from zone-melted rods of $Na_{0.75}CoO_2$, $Na_{0.78}CoO_2$ and $Na_{0.92}CoO_2$ grown using the floating zone method[25]. Single crystals from the rods were initially screened and a suitable piece in each case cut away.

Neutron Laue diffraction provides a powerful technique with which to probe atomic ordering in the bulk of a crystal sample. Data were collected on the single-crystal diffractometer (SXD) at the ISIS pulsed neutron source, Rutherford Appleton Laboratory, UK. SXD combines the white beam Laue technique with area detectors covering a solid-angle of 2π steradians, allowing comprehensive data sets to be collected. Samples were mounted on aluminium pins and cooled to 150 K using a closed-cycle helium refrigerator. A typical data set required four to five orientations to be collected for 12 to 35 hours per orientation, depending on sample size. Data were corrected for incident flux using a null-scattering V/Nb sphere. These data were then combined to a volume in reciprocal space and sliced to obtain individual planar cuts.

Additional measurements were performed using the flat-cone diffractometer E2 at the Hahn-Meitner Institute in Germany. E2 was used to determine accurate

modulation wavevectors. Measurements were performed using a variable-temperature cryostat down to a base temperature of 1.5 K and the temperature dependence was determined through the phase transition. Other small samples from the same rods were examined by synchrotron X-ray diffraction using the MAGS beamline at BESSY in Berlin, and electrical transport and magnetometry measurements were performed to assess their quality; see Supplementary Figs 3 and 5.

The structural superlattice neutron diffraction peaks are typically two orders of magnitude lower in intensity than the hexagonal Bragg peaks. The difficulty in observing the main magnetic Bragg reflections suggests that detection of any in-plane modulation of the magnetic structure would be difficult, even with polarized neutrons.

The quoted sodium concentrations x were determined using electron probe microanalysis. We note that neutron activation analysis gave a value of x roughly 5% lower. Small inclusions of a few per cent of the cobalt oxides (CoO and Co_3O_4) as well as Na_2O were found. The diffraction measurements show that these grow epitaxially on the host lattice and the impurity signal is straightforward to distinguish from that from the Na_xCoO_2.

Theory. The pair energy between Na^+ ions is the sum of a strongly repulsive short-range part preventing electronic shells from overlapping $V_{sr} = A\exp[-r/r_0] - C/r^6$ with $A = 424$ eV, $r_0 = 0.318$ Å and $C = 1.05$ eV Å^6 from the literature[26] and a long-range Coulomb contribution $V_c = e^2/4\pi\varepsilon\varepsilon_0 r$. At nearest- and next-nearest-neighbour distances, we obtain respectively $V_{sr}^1 = 2.53$ eV (considered infinite at room temperature) and $V_{sr}^2 = 0.06$ eV (we neglect V_{sr} at larger distances). The Coulomb landscape in Fig. 4 depends on one parameter: the effective dielectric constant ε. We take it to be isotropic and use the value $\varepsilon = 6$ to ensure stability of the long-range ordered phases up to room temperature. Our conclusions concerning vacancy clustering and the importance of the potential-well depth compared to hopping energy are quite robust and remain valid within a very large ε range.

This minimal model for the sodium system can be mapped onto a frustrated Ising model where a pseudo-spin state $I_i^z = +1/2$ denotes that the ith site of the honeycomb lattice is occupied, and $I_i^z = -1/2$ denotes that it is empty. The electrostatic and ionic shell interactions are represented by the hamiltonian:

$$H = \sum_{ij} \left(V_{ij} + c(r_{ij})\right)\left(I_i^z + 1/2\right)\left(I_j^z + 1/2\right) + \sum_i \left(-\mu + \Delta\left(1 + (-1)^i\right)\right)\left(I_i^z + 1/2\right)$$

where only shell repulsions V_{ij} for nearest-neighbour and next-nearest-neighbour interactions are treated as important; c absorbs the Coulomb repulsions. The chemical potential μ acts as a uniform field, and Na1 onsite energy Δ acts like a staggered field.

Long-range Coulomb interactions between all Co, O and Na ions on three-dimensional periodic lattices were summed with an accuracy of ten digits using a generalization of the Ewald method[27]. At finite temperature, Monte-Carlo simulations were performed in canonical and grand-canonical ensembles on periodic lattices including up to 1,800 Na sites. Vacancy clustering has been observed in the whole stability range $0 < T < 600$ K of Na_xCoO_2. For a variety of simple fractional fillings, long-range order of identical clusters appears at low temperature, and this melts into a glassy mixture of di-, tri- and quadri-vacancy clusters above $T \approx 400$ K. Full details will be reported elsewhere.

Received 21 July; accepted 13 December 2006.

1. Terasaki, I., Sasago, Y. & Uchinokura, K. Large thermoelectric power in $NaCo_2O_4$ single crystals. *Phys. Rev. B* **56,** R12685–R12687 (1997).
2. Lee, M. *et al.* Large enhancement of the thermopower in Na_xCoO_2 at high Na doping. *Nature Mater.* **5,** 537–540 (2006).
3. Takada, K. *et al.* Superconductivity in two-dimensional CoO_2 layers. *Nature* **422,** 53–55 (2003).
4. Delmas, C. *et al.* Electrochemical intercalation of sodium in Na_xCoO_2 bronzes. *Solid State Ionics* **3/4,** 165–169 (1981).
5. Zandbergen, H. W. *et al.* Sodium ion ordering in Na_xCoO_2: Electron diffraction study. *Phys. Rev. B* **70,** 024101 (2004).
6. Hubbard, J. Electron correlations in narrow energy bands. *Proc. R. Soc. Lond. A* **276,** 238–257 (1963).
7. Wang, Y. Y., Rogado, N. S., Cava, R. J. & Ong, N. P. Spin entropy as the likely source of enhanced thermopower in $NaxCo_2O_4$. *Nature* **423,** 425–428 (2003).
8. Amatucci, G. G., Tarascon, J. M. & Klein, L. C. CoO_2, the end member of the Li_xCoO_2 solid solution. *J. Electrochem. Soc.* **143,** 1114–1123 (1996).
9. Zhang, P., Capaz, R. B., Cohen, M. L. & Louie, S. G. Theory of sodium ordering in Na_xCoO_2. *Phys. Rev. B* **71,** 153102 (2005).
10. Huang, Q. *et al.* Low temperature phase transition and crystal structure of $Na_{0.5}CoO_2$. *J. Phys. Condens. Matter* **16,** 5803–5814 (2004).
11. Foo, M. L. *et al.* Charge ordering, commensurability, and metallicity in the phase diagram of the layered Na_xCoO_2. *Phys. Rev. Lett.* **92,** 247001 (2004).
12. de Vaulx, C. *et al.* Nonmagnetic insulator state in Na_1CoO_2 and phase separation of Na vacancies. *Phys. Rev. Lett.* **95,** 186405 (2005).
13. Huang, Q. *et al.* Coupling between electronic and structural degrees of freedom in the triangular lattice conductor Na_xCoO_2. *Phys. Rev. B* **70,** 184110 (2004).
14. Balsys, R. J. & Davis, R. L. Refinement of the structure of $Na_{0.74}CoO_2$ using neutron powder diffraction. *Solid State Ionics* **93,** 279–282 (1996).
15. Mukhamedshin, I. R. *et al.* ^{28}Na NMR evidence for charge order and anomalous magnetism in Na_xCoO_2. *Phys. Rev. Lett.* **93,** 167601 (2004).
16. Yokoi, M. *et al.* Magnetic correlation of Na_xCoO_2 and successive phase transitions of $Na_{0.5}CoO_2$—NMR and neutron diffraction studies. *J. Phys. Soc. Jpn* **74,** 3046–3056 (2005).
17. Hasan, M. Z. *et al.* Fermi surface and quasiparticle dynamics of $Na_{0.7}CoO_2$ investigated by angle-resolved photoemission spectroscopy. *Phys. Rev. Lett.* **92,** 246402 (2004).
18. Mukhamedshin, I. R., Alloul, H., Collin, G. & Blanchard, N. ^{59}Co NMR study of the Co states in superconducting and anhydrous cobaltates. *Phys. Rev. Lett.* **94,** 247602 (2005).
19. Bayrakci, S. P. *et al.* Magnetic ordering and spin waves in $Na_{0.82}CoO_2$. *Phys. Rev. Lett.* **94,** 157205 (2005).
20. Helme, L. M. *et al.* Three-dimensional spin fluctuations in $Na_{0.75}CoO_2$. *Phys. Rev. Lett.* **94,** 157206 (2005).
21. Slack, G. A. & Tsoukala, V. G. Some properties of semiconducting $IrSb_3$. *J. Appl. Phys.* **76,** 1665 (1994).
22. Nolas, G. S. *et al.* Semiconducting Ge clathrates: Promising candidates for thermoelectric applications. *Appl. Phys. Lett.* **73,** 178–180 (1998).
23. Nolas, G. S. *et al.* Effect of partial void filling on the lattice thermal conductivity of skutterudites. *Phys. Rev. B* **58,** 164–170 (1998).
24. Sakurai, H. *et al.* The role of the water molecules in novel superconductor, $Na_{0.35}CoO_2 \cdot 1.3H_2O$. *Physica C* **412–414,** 182–186 (2004).
25. Prabhakaran, D., Boothroyd, A. T., Coldea, R. & Charnley, N. R. Crystal growth of Na_xCoO_2 under different atmospheres. *J. Cryst. Growth* **271,** 74–80 (2004).
26. Fumi, F. G. & Tosi, M. P. Ionic sizes and born repulsive parameters in the NaCl-type alkali halides. *J. Phys. Chem. Solids* **25,** 31–52 (1964).
27. Sperb, R. An alternative to Ewald sums. *Mol. Simul.* **20,** 179–200 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Lee, J. Irvine, P. Radaelli, A. Daoud-Aladine, K. Kiefer, D. Argyriou, R. Coldea and M. Nohara for discussions, and T. Bowcock and A. Washbrook for the use of the MAP2 supercomputer at the University of Liverpool.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.R. (roger@drecam.saclay.cea.fr).

LETTERS

Tibetan plateau aridification linked to global cooling at the Eocene–Oligocene transition

Guillaume Dupont-Nivet[1], Wout Krijgsman[1], Cor G. Langereis[1], Hemmo A. Abels[2], Shuang Dai[3] & Xiaomin Fang[3,4]

Continental aridification and the intensification of the monsoons in Asia are generally attributed to uplift of the Tibetan plateau and to the land–sea redistributions associated with the continental collision of India and Asia[1–3], whereas some studies suggest that past changes in Asian environments are mainly governed by global climate[4–6]. The most dramatic climate event since the onset of the collision of India and Asia is the Eocene–Oligocene transition, an abrupt cooling step associated with the onset of glaciation in Antarctica 34 million years ago[7–9]. However, the influence of this global event on Asian environments is poorly understood. Here we use magnetostratigraphy and cyclostratigraphy to show that aridification, which is indicated by the disappearance of playa lake deposits in the northeastern Tibetan plateau, occurred precisely at the time of the Eocene–Oligocene transition. Our findings suggest that this global transition is linked to significant aridification and cooling in continental Asia recorded by palaeontological and palaeoenvironmental changes[10–12], and thus support the idea that global cooling is associated with the Eocene–Oligocene transition[13–15]. We show that, with sufficient age control on the sedimentary records, global climate can be distinguished from tectonism and recognized as a major contributor to continental Asian environments.

Global climate and regional tectonism affect palaeoenvironmental conditions that are recorded during the deposition and accumulation of sediments[16]. In Asia, climate models[1,3] have shown that uplift of the Tibetan plateau and retreat of the Paratethys epicontinental sea associated with the Indo-Asia collision during the Cenozoic era can result in monsoonal intensification, continental aridification and increased erosion. When recognized in the sedimentary records, these palaeoenvironmental conditions have thus been attributed to regional tectonism, although they can equally well indicate global climate changes[4–6]. The key to distinguishing climatic effects from those of tectonism is to date the regional expression of these processes in the geologic record with sufficient resolution and accuracy to enable correlation to global climatic events precisely calibrated in the marine realm. The exceptionally long (~30 million years, Myr) continuous sedimentary succession of the Xining basin[17] at the northeastern margin of the Tibetan plateau provides an excellent opportunity to study the sedimentary signature of tectonic and climatic processes during Eocene to Oligocene times. The stratigraphic sequence reveals a remarkably sharp and widespread change in depositional environment expressed regionally by the disappearance of massive gypsum layers. To determine the cause of this lithologic change, detailed lithofacies analyses and high-resolution dating were performed on two parallel sedimentary sections from the Xining basin (the Shuiwan and Xiejia sections; Fig. 1).

The lower part of the stratigraphy studied consists of regular alternations of laterally continuous gypsum layers and red mudstone beds (Fig. 2). The red mudstone intervals consist of silty clay virtually devoid of sedimentary structures, suggesting alteration during extended sub-aerial exposure after sheet-flood events supplied the clastic mud. They are typical of a distal alluvial fan environment[16]. Gypsiferous intervals are decimetre- to metre-thick tabular, nodular or laminar beds of alabastrine massive gypsum showing some chicken-wire structures and displacive enterolithic veins (Fig. 1). They can grade laterally into green mudstone with preserved lacustrine lamination and are indicative of various stages of chemical precipitation in a playa lake environment[16] (see Supplementary Table 1). The precipitation of evaporites in continental playa-type systems is

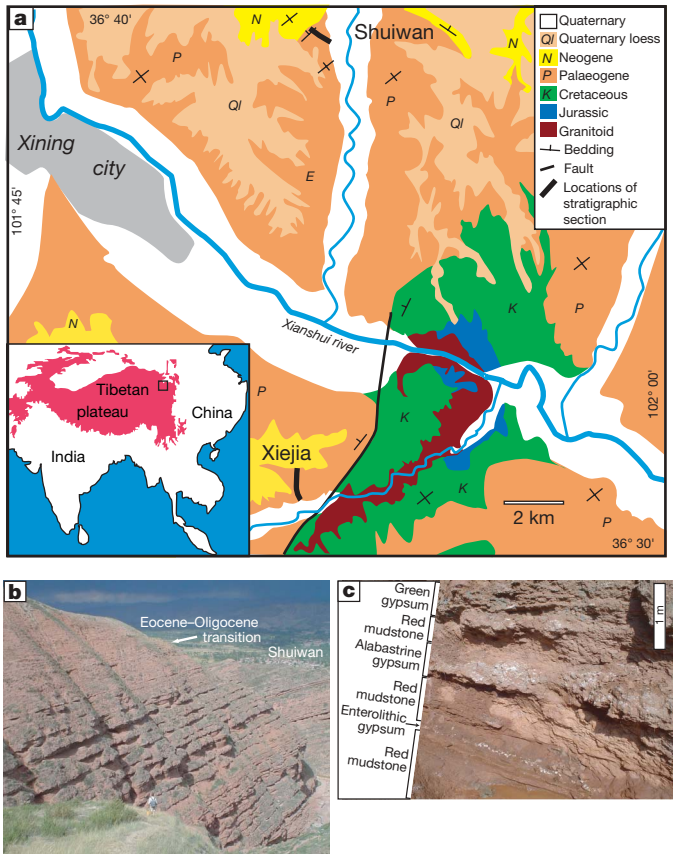


Figure 1 | Geologic setting. a, Location of studied stratigraphic sections. **b**, Field view to northeast, showing part of the Shuiwan section, including the end of cyclic gypsiferous deposition at the Eocene–Oligocene transition. **c**, Typical gypsum–mudstone cyclic alternations displaying characteristic lithofacies described in the text.

[1]Paleomagnetic Laboratory 'Fort Hoofddijk', Department of Earth Sciences, Utrecht University, Budapestlaan 17, [2]Stratigraphy-Paleontology, Department of Earth Sciences, Utrecht University, Budapestlaan 4, 3584 CD Utrecht, The Netherlands. [3]Lanzhou University, Key Laboratory of Western China's Environmental Systems, Gansu 730000, China. [4]Institute of Tibetan Plateau Research, Chinese Academy of Science, P O Box 2871, Beijing 100085, China.

chiefly determined by the supply of drained atmospheric water, ultimately providing solutes through groundwater inflow and overland runoff[18]. Gypsum intercalations within the red mudstones are thus indicative of periods with significantly higher water supply providing sufficient solutes to accumulate metre-scale gypsum layers[18]. The regularity of the gypsum/mudstone alternations indicate periodical fluctuations in atmospheric water supply to the basin, suggesting a cyclic mechanism most probably controlled by climate oscillations.

In the upper part of the stratigraphy, fully developed gypsum layers disappear. Sediments are dominated by homogenous red mudstones with minor interstitial gypsum content and occasional sandstone layers indicative of a distal alluvial fan depositional environment, without playa lake intercalations. The inhibition of gypsum accumulation between the lower and the upper part of the stratigraphy implies a permanent drop in the supply of atmospheric water that resulted in pronounced aridification of the Xining basin.

Consistent with this, regional palaeoenvironmental records also document aridification in the Eocene to Oligocene interval considered here. The fossil pollen records from the studied sediments of the Xining basin[19] indicate a drying and cooling environment, as indicated by an increase in herbs and conifers. North of the Tibetan plateau, an abrupt faunal turnover is linked to cooling and aridification, as indicated by the evolution of mammal tooth patterns and a drastic size decrease from large Eocene species (including some of the largest terrestrial mammals ever known) to small Oligocene species dominated by rodents[10]. Moreover, the first extensive set of $\delta^{18}O$ and $\delta^{13}C$ isotopic data reveals a positive shift between Eocene and Oligocene continental strata that has been attributed to aridification[12].

To determine whether this aridification can be linked to tectonism or known climatic events in the Eocene to Oligocene interval, we accurately dated the sediments. Magnetostratigraphic sampling of the two sections was performed to determine the pattern of normal and reversed polarity intervals of the Earth's palaeomagnetic field. Palaeomagnetic analysis (see Methods and Supplementary Information) and the excellent consistency between the two records indicate a primary origin of the rock magnetization and the reliability of polarity intervals (Fig. 2). Our results improve and corroborate the long magnetostratigraphic records previously obtained from the Xining basin[17], showing a distinctive pattern of two long reversed intervals (R1 and R2) separated by a shorter normal interval (N2). This pattern provides a unique correlation to chrons C12r and C13r of the geomagnetic polarity timescale (GPTS[20]), which indicates that the sampled interval straddles the Eocene–Oligocene boundary, in agreement with palaeontological constraints[17,19] (see Methods). According to the GPTS correlation, the average period of gypsum–mudstone alternations is 109 thousand years (109 kyr) at Xiejia and 93 kyr at Shuiwan. These values are consistent with the ~100-kyr periodicity of the Earth's orbital eccentricity, suggesting that the supply of atmospheric water, periodically sufficient for gypsum

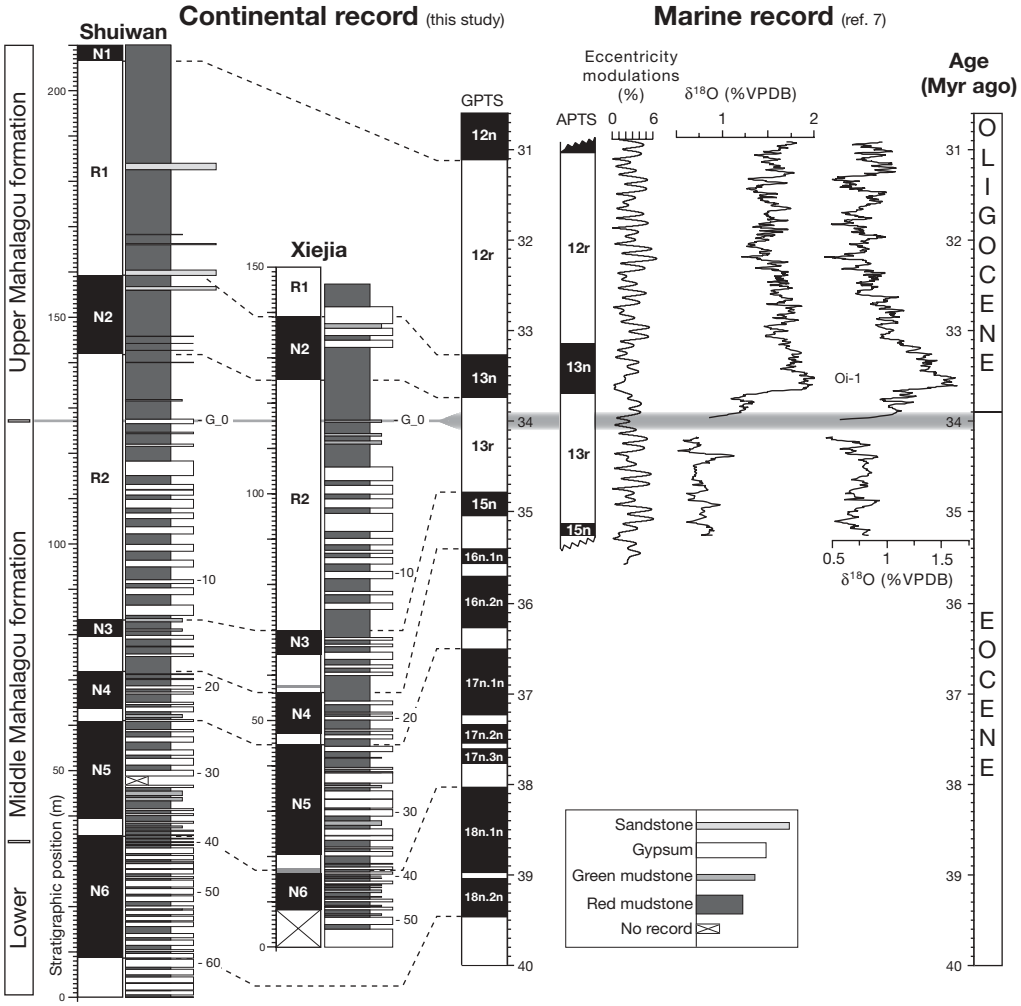


Figure 2 | Stratigraphic correlations. Lithostratigraphic descriptions are given in metres from the bottom of the Shuiwan and Xiejia sections. Eccentricity-driven gypsum–mudstone cycles are numbered down from the highest correlatable level G_0 at the lithofacies change. Observed normal (N) and reverse (R) polarity zones (N1 to N6) are correlated to chrons of the GPTS[20] and the APTS[7]. Eccentricity modulation model is from ref. 22. The stable isotope shift is from the marine record at the Eocene–Oligocene transition[7]. VPDB, Vienna Pee-Dee Belemnite standard.

deposition, was controlled by eccentricity-driven climate variability. Pronounced aridification is defined by the interruption of this periodic system of basin-wide gypsum generation, above the highest gypsiferous level correlatable between the two sections (G_0 on Fig. 2). This layer marks the end of the distinctive pattern of regular gypsum–mudstone cycles, followed by mudstone-dominated successions with occasional non-correlative sandy and gypsiferous levels. In both sections, G_0 occurs near the top of C13r. Using linear interpolation between GPTS chron boundaries, G_0 is estimated at 34 Myr ago (33.9 Myr ago at Xiejia and 34.0 Myr ago at Shuiwan, Supplementary Table 4).

This age corresponds well with the most important step in Cenozoic global cooling 34 Myr ago, characterized by the onset of permanent Antarctic ice sheets at the Eocene–Oligocene transition and marked in the marine records by an abrupt increase in $\delta^{18}O$ and $\delta^{13}C$ values (the Oi-1 event)[7–9,21] (Fig. 2). This marine isotope increase was recently precisely documented as a two-step shift (40 kyr each) separated by an intermediate plateau (200 kyr)[7]. To formally compare our results with this marine record, we calibrate our data to the astronomically tuned polarity timescale (APTS) of this study[7], providing chron boundaries dated with respect to the variation of Earth's orbital parameters[22]. Linear interpolation between these chron boundaries also indicates a 34-Myr age for G_0 (33.9 at Xiejia and 34.1 at Shuiwan; Supplementary Table 4). Furthermore, the recognized eccentricity forcing of the gypsum–mudstone cyclicity enables us to test the consistency of our records directly with the astronomically tuned marine record. The number of eccentricity cycles expected to separate G_0 from the underlying chron C15n—according to the APTS—is indeed consistent at both sections with the number of observed gypsum–mudstone alternations separating G_0 from N3 (13 alternations at Shuiwan and Xiejia). This result independently validates our age estimate for G_0 at 34 Myr ago, indicating that the start of aridification after G_0 is coeval with the marine isotope shift of the climate transition (Fig. 2).

This remarkable correlation provides the first direct link between the Eocene–Oligocene climate transition and the aridification of continental Asia. In view of this finding, existing Eocene–Oligocene records of Asian environments must be interpreted with respect to climate change. The positive $\delta^{18}O$ and $\delta^{13}C$ shifts recorded north of the Tibetan plateau[12] can be attributed to the combined effects of aridification and the positive shift in isotopic values of ocean water (the ultimate source of meteoric water) at the Eocene–Oligocene transition. Abrupt changes in the Asian fauna in response to cooling and aridification (the Mongolian Remodelling)[10] can now be confidently linked to the Eocene–Oligocene transition and correlated to the large mammalian turnover in Europe (the 'Grande Coupure'), as well as globally reported continental and marine records of comparable faunal and floral turnovers at this time[13,21]. This supports the paradigm that the Eocene–Oligocene transition is associated with global atmospheric cooling and contrasts with recent claims, based on a few localized continental palaeotemperature estimates, that it may be linked to changes in Antarctic ice volume only[14,15]. Records of cooling and aridification[10–12,19] in the Asian interior are thus attributed to cooling of global ocean temperatures, reducing moisture supply to continental interiors. This effect may have been enhanced by a large Paratethys retreat induced by glacio-eustatic sea level lowering driven by Antarctic glaciation. Lastly, increasing sediment accumulation recorded over the East and Southeast Asian offshore region at about 33 Myr ago[23] now suggests that Eocene–Oligocene global cooling was also associated with increased regional erosion and sediment transport to the sea.

Our study demonstrates that global climate cooling must be recognized as a major contributor to Asian palaeoenvironment. It substantiates previous propositions that Pliocene climate deterioration or late Miocene global cooling essentially governed Asian depositional environments[4–6]. The growing evidence that global cooling is associated with continental aridification along with intensification of monsoons and increased regional erosion remains to be tested by climate models. General circulation models have mainly focused on the effect of regional tectonism on Asian and global climate[1,3,24], showing that Tibetan uplift and Paratethys retreat can also increase Asian continental aridity, monsoons and erosion. Ultimately, Tibetan uplift can lead to global cooling through atmospheric circulation perturbations and carbon dioxide lowering associated with increased rock weathering[25].

Our results do not exclude that, in conjunction with the Eocene–Oligocene transition, tectonic processes have also contributed to Asian palaeoclimatic conditions. However, to establish the importance of these processes in driving Asian or global climate at this time, improved age control on their sedimentary expression is required. The Paratethys retreat is bracketed between late Eocene and latest Miocene time[26] and estimates for Tibetan plateau uplift range from the early Cenozoic at least 35 Myr ago[27,28] to the Pliocene 2–4 Myr ago[1]. In time, accurate and precise dating of Tibetan uplift may finally confirm the hypothesis that it was the primary cause of the Eocene–Oligocene climate transition through global atmospheric carbon dioxide lowering, instead of the recently challenged paradigm that the Eocene–Oligocene transition was triggered by the opening of a sea passage around Antartica[7–9,29]. Our study illustrates that distinguishing climate effects from tectonism can be accomplished by high-resolution dating, using magnetostratigraphy validated independently by the study of orbitally driven cyclicity in sediments[30]. This seems a promising tool for Asian sedimentary records, which often show cyclicity but are devoid of volcanics for radiometric calibration.

METHODS

Palaeomagnetic analysis. Palaeomagnetic sampling was performed in the course of two field seasons using a standard portable coring device; 187 and 110 sampling levels were collected in the 209.5-m-thick Shuiwan section and the 145.8-m-thick Xiejia section respectively (average sampling interval of 1.2 m). Palaeomagnetic analysis yielded reliable primary characteristic remanent magnetization directions at 206 levels (see Supplementary Figures 1 and 2 and Supplementary Tables 2 and 3). Polarity zones at the Shuiwan (12 zones) and Xiejia (10 zones) sections are defined by palaeomagnetic directions from three or more levels (Supplementary Fig. 3). Two questionable short polarity zones are defined by one sample only (depicted in grey on Fig. 2). Apart from these short intervals, the excellent consistency between the two magnetostratigraphic records strongly confirms the primary origin of the rock magnetization and the reliability of polarity intervals.

Palaeontological constraints. Palaeontological constraints provide a starting point for our age model (see palaeontological references in ref. 17). A relatively rich record of fossil mammals, ostracods, charophytes and pollen have been collected at the location of the two sampled sections (Xiejia and Shuiwan) including the Mahalagou formation (of interest for this study), the underlying Honggou formation, and the overlying Xiejia formation. The upper part of the Xiejia formation contains a well-documented mammal assemblage (the Xiejia fauna) that is described to be from the lowermost Miocene age (>20 Myr ago), implying that most of the Xiejia formation below this fossil locality is of Oligocene age. In addition, the lower part of the Xiejia formation is ascribed a late Oligocene age, the Mahalagou formation an early Oligocene age and the Honggou formation an Eocene age, on the basis of ostracods, charophytes and (mostly) pollen records[19]. Taken together, existing palaeontological constraints point to an early Oligocene age assignment for the sampled strata in the Mahalagou formation.

Magnetostratigraphic correlation. The complete pattern of magnetic polarity zones from the Honggou, Mahalagou and Xiejia formations recorded at the Xiejia and Shuiwan sections is provided by ref. 17. From this long record, two long reversed polarity zones separated by a shorter normal zone stand out clearly within the Mahalagou formation. Our focused palaeomagnetic sampling and analysis of this characteristic interval confirms this pattern and excludes the occurrences of the questionable short normal polarity intervals reported in the Xiejia section[17]. Above and below the pair of long reversed polarity zones are numerous shorter zones. When compared to the GPTS, the Early Oligocene to Late Eocene time interval is characterized by two long reversed chrons (C12r and C13r) separated by a shorter normal chron (C13n). Above and below this pair of long reversed chrons, the GPTS indicates shorter chrons. Given the palaeontological constraints and the pattern of magnetic polarity zones, the simplest

interpretation is to correlate the pair of long reversed polarity zones of the Mahalagou formation to the pair of reversed polarity chrons (C12r and C13r).

Although the shorter length of R1 relative to R2 is inverse to the pattern of C12r–C13r in the GPTS, alternative correlations of this pair of long reversed intervals are practically impossible. Correlating the long observed reversed intervals higher on the GPTS where the reversal rate is high would imply, at least, an order-of-magnitude increase in accumulation rates, fortuitously during the duration of these chrons. Correlation below would require these long reversed chrons to be expressed higher up in the section, where recorded polarity zones are short, ultimately implying at least an order-of-magnitude decrease in accumulation rates. Such large variations would probably be indicated by facies variations (for example, conglomerates, disconformity or unconformity) that are not observed.

Below R2, we correlate N3 to C15n, N4 to C16n.1n–C16.2n, N5 to C17n.1n–C17n.2n–C17n.3n and N6 to C18n.1n–C18n.2n (Fig. 2). Supporting these correlations are the resulting consistent sedimentation rates calculated by linear interpolation between these chron boundaries (Supplementary Table 4). However, this correlation implies that four short reversed intervals have not been detected: C16n.1r = 159 kyr; C17n.1r = 131 kyr; C17n.2r = 72 kyr; and C18n.1r = 79 kyr. Given the low sedimentation rates, we argue that the following can explain the absence of these short intervals: (1) insufficient sampling resolution (average 1.2 m); (2) the absence of sampling in gypsum intervals yielding no palaeomagnetic signal; and (3) possible gaps in such sub-aerial deposits. An alternative in which no chrons are missed may be considered by correlating N5 to C16n.2n and N6 to C17n.1n. This alternative correlation would not change the magnetostratigraphic correlation of G_0 within C13r but we discarded it because the expected long normal interval C18n is not observed below N6 in the existing record[17], which shows instead a long reversed interval.

Received 25 July; accepted 5 December 2006.

1. An, Z., Kutzbach, J. E., Prell, W. L. & Porter, S. C. Evolution of Asian monsoons and phased uplift of the Himalaya-Tibetan plateau since Late Miocene times. *Nature* **411,** 62–66 (2001).
2. Guo, Z. T. *et al.* Onset of Asian desertification by 22 Myr ago inferred from loess deposits in China. *Nature* **416,** 159–163 (2002).
3. Ramstein, G., Fluteau, F., Besse, J. & Joussaume, S. Effect of orogeny, plate motion and land–sea distribution on Eurasian climate change over the past 30 million years. *Nature* **386,** 788–795 (1997).
4. Gupta, A. K., Singh, R. K., Joseph, S. & Thomas, E. Indian Ocean high-productivity event (10–8 Ma): Linked to global cooling or to the initiation of the Indian monsoons? *Geology* **32,** 753–756 (2004).
5. Molnar, P. Late Cenozoic increase in accumulation rates of terrestrial sediments: How might climate change have affected erosion rates? *Annu. Rev. Earth Planet. Sci.* **32,** 67–89 (2004).
6. Zhang, P., Molnar, P. & Downs, W. Increased sedimentation rates and grain sizes 2–4 Myr ago due to the influence of climate change on erosion rates. *Nature* **410,** 891–897 (2001).
7. Coxall, H. K., Wilson, P. A., Palike, H., Lear, C. H. & Backman, J. Rapid stepwise onset of Antarctic glaciation and deeper calcite compensation in the Pacific Ocean. *Nature* **433,** 53–57 (2005).
8. DeConto, R. M. & Pollard, D. Rapid Cenozoic glaciation of Antarctica induced by declining atmospheric CO2. *Nature* **421,** 245–249 (2003).
9. Zachos, J. C. & Kump, L. R. Carbon cycle feedbacks and the initiation of Antarctic glaciation in the earliest Oligocene. *Glob. Planet. Change* **47,** 51–66 (2005).
10. Meng, J. & McKenna, M. C. Faunal turnovers of Palaeogene mammals from the Mongolian Plateau. *Nature* **394,** 364–367 (1998).
11. Garzione, C., Ikari, M. J. & Basu, A. R. Source of Oligocene to Pliocene sedimentary rocks in the Linxia basin in northeastern Tibet from Nd isotopes: Implications for tectonic forcing of climate. *Geol. Soc. Am. Bull.* **117,** 1156–1166, doi:10.1130/B25743.1 (2005).
12. Graham, S. A. *et al.* Stable isotope records of Cenozoic climate and topography, Tibetan plateau and Tarim basin. *Am. J. Sci.* **305,** 101–118 (2005).
13. Ivany, L. C., Patterson, W. P. & Kyger, C. L. Cooler winters as a possible cause of mass extinctions at the Eocene/Oligocene boundary. *Nature* **407,** 887–890 (2000).
14. Grimes, S. T., Hooker, J. J., Collinson, M. E. & Mattey, D. P. Summer temperatures of late Eocene to early Oligocene freshwaters. *Geology* **33,** 189–192 (2005).
15. Kohn, M. J. *et al.* Climate stability across the Eocene-Oligocene transition, southern Argentina. *Geology* **32,** 621–624 (2004).
16. Talbot, M. R. & Allen, P. A. in *Sedimentary Environments: Processes, Facies and Stratigraphy* (ed. Reading, H. G.) 83–124 (Blackwell Publishing Inc., Oxford, 1996).
17. Dai, S. *et al.* Magnetostratigraphy of Cenozoic sediments from the Xining Basin: Tectonic implications for the northeastern Tibetan Plateau. *J. Geophys. Res.* **111,** doi:10.129/2005JB004187 (2006).
18. Yechieli, Y. & Wood, W. W. Hydrogeologic processes in saline systems: playas, sabkhas and saline lakes. *Earth Sci. Rev.* **58,** 343–365 (2002).
19. Wang, D.-N., Sun, X.-Y. & Zhao, Y.-N. Late Cretaceous to Tertiary palynofloras in Xinjiang and Qinghai, China. *Rev. Palaeobot. Palynol.* **65,** 95–104 (1990).
20. Ogg, J. G. & Smith, A. G. in *A Geologic Time Scale 2004* (eds Gradstein, F. M., Ogg, J. G. & Smith, A. G.) 63–86 (Cambridge Univ. Press, Cambridge, UK, 2004).
21. Prothero, D. R. in *Encyclopedia of Geology* (eds Selley, R., Cocks, R. & Plimer, I.) 472–478 (Elsevier, London, 2004).
22. Laskar, J. *et al.* A long-term numerical solution for the insolation quantities of the Earth. *Astron. Astrophys.*, doi:10.1051/0004–6361:20041335 (2004).
23. Clift, P. D., Layne, G. D. & Blusztajn, J. Marine sedimentary evidence for monsoon strengthening, Tibetan uplift and drainage evolution in Asia. Continent-ocean interactions in the East Asian marginal seas. *AGU Monogr.* **149,** 255–282 (2004).
24. Kutzbach, J. E. & Behling, P. Comparison of simulated changes of climate in Asia for two scenarios: Early Miocene to present, and present to future enhanced greenhouse. *Glob. Planet. Change* **41,** 157–165 (2004).
25. Ruddiman, W. F., Kutzbach, J. E. & Prentice, I. C. in *Tectonic Uplift and Climate Change* (ed. Ruddiman, W. F.) 203–235 (Plenum, New York, 1997).
26. Popov, S. *et al.* Lithological-paleogeographic maps of Paratethys 10 maps Late Eocene to Pliocene. *Courier Forschungsinstitut Senckenberg* **250,** 1–42 (2004).
27. Rowley, D. B. & Currie, B. S. Palaeo-altimetry of the late Eocene to Miocene Lunpola basin, central Tibet. *Nature* **439,** 677–681 (2006).
28. Harris, N. B. W. The elevation history of the Tibetan Plateau and its implications for the Asian monsoon. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **241,** 4–15 (2006).
29. Huber, M. *et al.* Eocene circulation of the Southern Ocean: Was Antarctica kept warm by subtropical waters? *Paleoceanography* **19,** doi:10.1029/2004PA001014 (2004).
30. Krijgsman, W. *et al.* Revised astrochronology for the Ain el Beida section (Atlantic Morocco): No glacio-eustatic control for the onset of the Messinian Salinity Crisis. *Stratigraphy* **1,** 87–101 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Reviews by C. Garzione and discussions with F. Hilgen and M. Szurlies greatly improved the original manuscript. We thank H. Pälike and P. Wilson for sharing chron age calibrations. Lanzhou University students Dang Y. and Meng Q. provided logistical assistance. This project was funded through a 'Marie Curie' Fellowship from the European Union and a 'Veni' grant from the Netherlands science foundation (NWO) to G.D.-N.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to G.D.-N. (gdn@geo.uu.nl).

LETTERS

Large temperature drop across the Eocene–Oligocene transition in central North America

Alessandro Zanazzi[1], Matthew J. Kohn[1], Bruce J. MacFadden[2] & Dennis O. Terry Jr[3]

The Eocene–Oligocene transition towards a cool climate (~33.5 million years ago) was one of the most pronounced climate events during the Cenozoic era[1]. The marine record of this transition has been extensively studied. However, significantly less research has focused on continental climate change at the time, yielding partly inconsistent results on the magnitude and timing of the changes[2–8]. Here we use a combination of *in vivo* stable isotope compositions of fossil tooth enamel with diagenetic stable isotope compositions of fossil bone to derive a high-resolution (about 40,000 years) continental temperature record for the Eocene–Oligocene transition. We find a large drop in mean annual temperature of 8.2 ± 3.1 °C over about 400,000 years, the possibility of a small increase in temperature seasonality, and no resolvable change in aridity across the transition. The large change in mean annual temperature, exceeding changes in sea surface temperatures at comparable latitudes[9,10] and possibly delayed in time with respect to marine changes by up to 400,000 years, explains the faunal turnover for gastropods, amphibians and reptiles, whereas most mammals in the region were unaffected. Our results are in agreement with modelling studies that attribute the climate cooling at the Eocene–Oligocene transition to a significant drop in atmospheric carbon dioxide concentrations.

During the Eocene–Oligocene transition (EOT), at ~33.5 Myr ago, climate rapidly changed, marking a dramatic move from a 'greenhouse' to an 'icehouse' world[1]. Likely causes for this shift include opening of marine gateways[11], and/or reduction of atmospheric CO_2 concentrations[12]. The marine record of the EOT has been extensively studied, and shows major faunal turnovers that affected numerous organisms on a global scale[13–15]. In addition, oxygen isotope records of benthic foraminifera show a rapid increase in $\delta^{18}O$ near the EOT, due to combined ice sheet growth on Antarctica and deep-water cooling[1]. The increase in $\delta^{18}O$ of the EOT was accomplished in two 40-kyr-long steps separated by a 200-kyr-long intermediate plateau[16].

In contrast to the marine record of the EOT, significantly less research has focused on the continental climate change during this time. Terrestrial faunal records in North America indicate major extinctions for gastropods, amphibians and reptiles[2,8], but little change to mammals, except for brontotheres[3]. Unfortunately, North American continental climate data are ambiguous, implying changes in mean annual temperature (MAT) ranging from 0 to 8 °C (refs 2, 4, 5). The timing relationship between the continental and oceanic climate change of the EOT also remains unclear, with either climatic synchrony[7,8] or a possible continental time lag of 200–800 kyr (refs 3, 6). These inconsistencies, along with the coarse time resolution of these studies (~1–2 Myr time slices), preclude accurate understanding of links between continental and oceanic climate during major episodes of global change.

In this study, we analysed the stable isotope composition of fossil bones and teeth collected from the late Eocene–early Oligocene rocks in northwestern Nebraska, southwestern South Dakota, and eastern Wyoming (Fig. 1). The northern Great Plains of the US represents an ideal study site by virtue of its unparalleled fossil record and chronological controls. Tephrochronologic, single crystal $^{40}Ar/^{39}Ar$ dates on volcanic ashes[17], and biostratigraphic[18] and magnetostratigraphic[19] studies indicate sufficiently high sedimentation rates to allow high-resolution relative chronologic control (±a few tens of kyr) comparable to marine studies.

Bones and teeth in living animals are composed of calcium phosphate with significant substitutions of CO_3^{2-}. Stable isotope systematics in biogenic phosphates have been reviewed in ref. 20. Tooth enamel is resistant to isotopic exchange, so after burial and fossilization it retains an *in vivo* signal, largely reflecting dietary ($\delta^{13}C$) and, in obligate drinkers, local water ($\delta^{18}O$) compositions. Seasonal climate changes are preserved in $\delta^{18}O$ zoning along the length of each tooth. Mammals are homeotherms, so there is no temperature dependence to these isotope compositions. In contrast, the CO_3^{2-} component of bone is altered on timescales of 20–50 kyr, resetting with an important temperature-dependence to diagenetic compositions comparable to palaeosols[21]. $\delta^{18}O$ values of CO_3^{2-} in fossil bone are conceptually similar to foraminiferal CO_3^{2-} in that they reflect both local water composition and temperature. $\delta^{13}C$ values in bone reflect $\delta^{13}C$ of soil CO_2, which depends on plant organic matter composition and penetration of atmospheric CO_2 (ref. 22). Because C_4 plants were not yet abundant in the late Eocene and early Oligocene[23], $\delta^{13}C$ values principally reflect C_3 water stress, plant productivity, and canopy structure, increasing with increasing aridity. In this study, we use $\delta^{13}C$

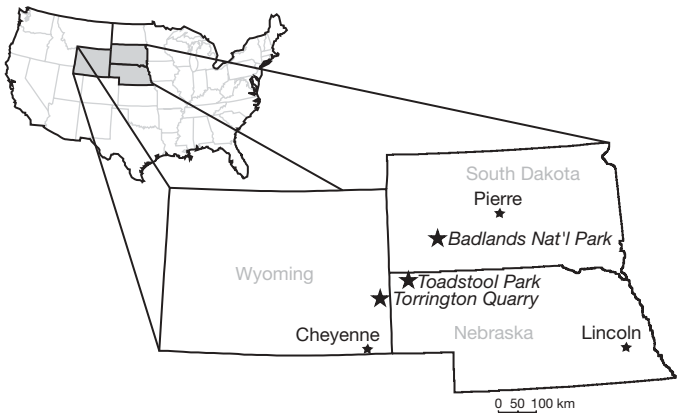


Figure 1 | Location map showing the sampling sites. Toadstool Park (northwestern Nebraska), Badlands National Park (southwestern South Dakota), and Torrington Quarry (eastern Wyoming).

[1]Department of Geological Sciences, University of South Carolina, Columbia, South Carolina 29208, USA. [2]Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611, USA. [3]Department of Geology, Temple University, Philadelphia, Pennsylvania 19122, USA.

values of bone to monitor changes in aridity, and combined $\delta^{18}O$ values of tooth enamel and bone to distinguish changes in water composition from changes in temperature.

The analysed samples include 231 bone fragments from various (usually unknown) mammalian taxa and turtles, and third molars (M3) of four of the most abundant White River mammals: *Mesohippus* (an ancestral horse), *Merycoidodon* (a sheep-sized artiodactyl), *Leptomeryx* (a small artiodactyl) and *Subhyracodon* (an ancestral rhino), which were all present before and after the transition. We also analysed middle molars (M2) and premolars for *Mesohippus* and premolars for *Subhyracodon* to ensure that we sampled all seasons. Oxygen isotope analyses were performed on the CO_3^{2-} component of tooth enamel and, for a few selected samples, also on the PO_4^{3-} component to test for diagenetic alteration. The *in vivo* relationship between $\delta^{18}O_{CO3}$ and $\delta^{18}O_{PO4}$ obtained for these samples indicates no significant diagenetic isotopic alteration of tooth enamel (see Supplementary Information). Isotope zoning in *Subhyracodon* and *Mesohippus* suggests that all seasons are sampled.

For tooth enamel $\delta^{18}O$ (Fig. 2), statistical tests (two-tailed *t*-tests unless mentioned below) indicate that mean values of Eocene versus Oligocene teeth cannot be distinguished for the following four mammals (*n* values are shown in the order Eocene then Oligocene): *Mesohippus* ($P = 0.436$; $n = 29$, 91), *Merycoidodon* ($P = 0.664$; $n = 19$, 11), *Leptomeryx* ($P = 0.605$; $n = 15$, 7) and *Subhyracodon* ($P = 0.607$, two-tailed Mann–Whitney test; $n = 89$, 30). Statistical tests also indicate that mean $\delta^{18}O$ values among the taxa are not statistically different, except *Merycoidodon*, which has a lower $\delta^{18}O$ ($P < 0.001$, single factor analysis of variance and Tukey test), and that $\delta^{18}O$ variances for Eocene versus Oligocene teeth cannot be distinguished (by *F*-test) for *Mesohippus* ($P = 0.200$), *Merycoidodon* ($P = 0.479$), *Leptomeryx* ($P = 0.856$) and *Subhyracodon* ($P = 0.672$).

Bone carbonate $\delta^{13}C$ values (Fig. 3) show no resolvable change across the EOT: -7.77 ± 0.14‰ (± 2 standard errors, s.e.) versus -7.62 ± 0.13‰ (± 2 s.e.). In contrast, $\delta^{18}O$ values show a ~1.7‰ positive shift: 22.98 ± 0.18‰ (± 2 s.e.) versus 24.63 ± 0.22‰ (± 2 s.e.). Linear regressions of $\delta^{18}O$ versus time resolve no statistically significant trends at 95% confidence for either the Eocene or Oligocene data subsets. The isotope shift in our continental record appears to lag the marine transition by ~400 kyr, but this may simply reflect sparse data in the time range 33.7–33.3 Myr ago. Similarly, our data do not sample the largest shift to marine $\delta^{13}C$ values at 33.4–33.6 Myr ago.

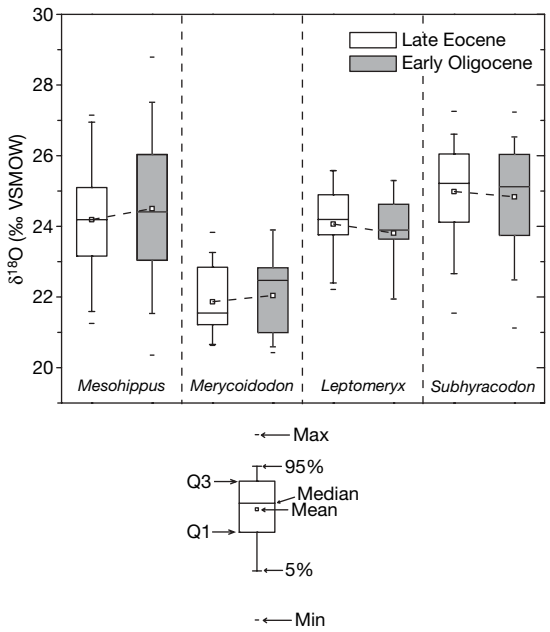


Figure 2 | 'Box and whiskers' plots for enamel $\delta^{18}O$. Lower and upper sides of boxes indicate lower and upper quartiles, respectively. Inside boxes, solid lines indicate median and squares indicate mean. Length of whiskers indicates 5th to 95th percentile range. Outside boxes, small tick-marks indicate minimum and maximum values. All taxa show insignificant changes in mean and median enamel $\delta^{18}O$ across the EOT, suggesting no change in rainwater composition. *Mesohippus*, *Merycoidodon* and *Subhyracodon* show a slightly higher Oligocene versus Eocene range in $\delta^{18}O$, suggesting a slightly higher Oligocene versus Eocene MART.

For the bone $\delta^{13}C$ data, the lack of a resolvable change across the EOT implies that major changes in water stress or canopy structure did not occur, although minor changes might have occurred without affecting $\delta^{13}C$. For the enamel $\delta^{18}O$ data, it is important that modern horses and rhinos are quite water-dependent because of their digestive physiology[24], which is thought to have developed by the early Eocene[25]. Therefore, the similar or lower mean $\delta^{18}O$ for all taxa imply that they were all water-dependent, even the artiodactyls. As a consequence, the lack of a resolvable change to mean enamel $\delta^{18}O$ across the EOT implies that water composition remained unchanged. In addition, the indistinguishable variance of Eocene versus Oligocene enamel $\delta^{18}O$ suggests no major change in temperature seasonality (mean annual range of temperature, MART) across the transition, although the small increase in the range of *Mesohippus*, *Merycoidodon* and *Subhyracodon* enamel $\delta^{18}O$ may reflect a slightly higher Oligocene versus Eocene MART.

Assuming that digestive physiology conferred similar dependencies of oxygen isotopes on climate variables for fossil equids and rhinos as observed in modern horses, we used the oxygen isotope data from *Mesohippus* to calculate MAT, changes in MAT (ΔMAT), and a minimum estimate of MART: MAT(Eocene) = 21.0 ± 10.1 °C, MAT(Oligocene) = 13.1 ± 9.5 °C; ΔMAT(EOT) = -8.2 ± 3.1 °C; and MART(Eocene) $\geq 21.9 \pm 7.5$ °C, MART(Oligocene) $\geq 25.8 \pm 8.8$ °C (all errors ± 2 s.e.; see Supplementary Information). Our MAT estimates are broadly compatible with those from faunal change (~20 °C and ~15 °C for the Eocene and Oligocene[2]), and leaf morphology (~19 °C and ~12 °C; ref. 26) studies. The decrease in MAT of ~8 °C is far larger than found in marine records, which support at most 2–3 °C of cooling at low to mid-latitudes[9,10]. Oligocene MART estimates agree with those from leaf morphology (~22 °C and 20–30 °C; refs 5, 26). However, our Eocene MART estimates are larger than those from these previous investigations (~10 °C and ~15 °C; refs 5, 26). We find no evidence for a systematic shift in cold-month mean temperatures, as observed in fish otoliths from the Gulf of Mexico[27]. Within uncertainty, the ~8 °C temperature drop could have caused cold-month mean temperatures in the Oligocene to approach freezing, probably accounting for the demise of many terrestrial heterotherms[2,8], while almost all large mammals survived[3].

Because ocean water $\delta^{18}O$ probably increased by ~0.75‰ (ref. 16), constant meteoric water compositions imply a compensating, about −0.75‰, temperature-induced decrease in $\delta^{18}O$. That is, the temperature coefficient for meteoric water $\delta^{18}O$ associated with the approximately −8 °C change in MAT is ~0.1‰ per °C, consistent with theoretical models (albeit for coastal areas) that suggest minimal temperature coefficients at mid-latitudes[28].

The large temperature drop in the North American mid-continent during the EOT supports the results of recent simulations that attribute the climatic deterioration during this time to a major decrease in the atmospheric partial pressure of CO_2 (ref. 12). The opening of Southern Ocean gateways probably had a minor role in driving the transition because the Antarctic Circumpolar Current is expected to produce warming of middle and high latitudes of the Northern Hemisphere[29], directly contrary to our results.

In sum, enamel and bone isotope compositions indicate a large MAT decrease, a possible small MART increase, and no resolvable change in aridity across the transition. Therefore, in contrast to previous studies[2,8,27], our results indicate that a decrease in MAT, rather than an increase in seasonality or aridity, is the primary cause of the

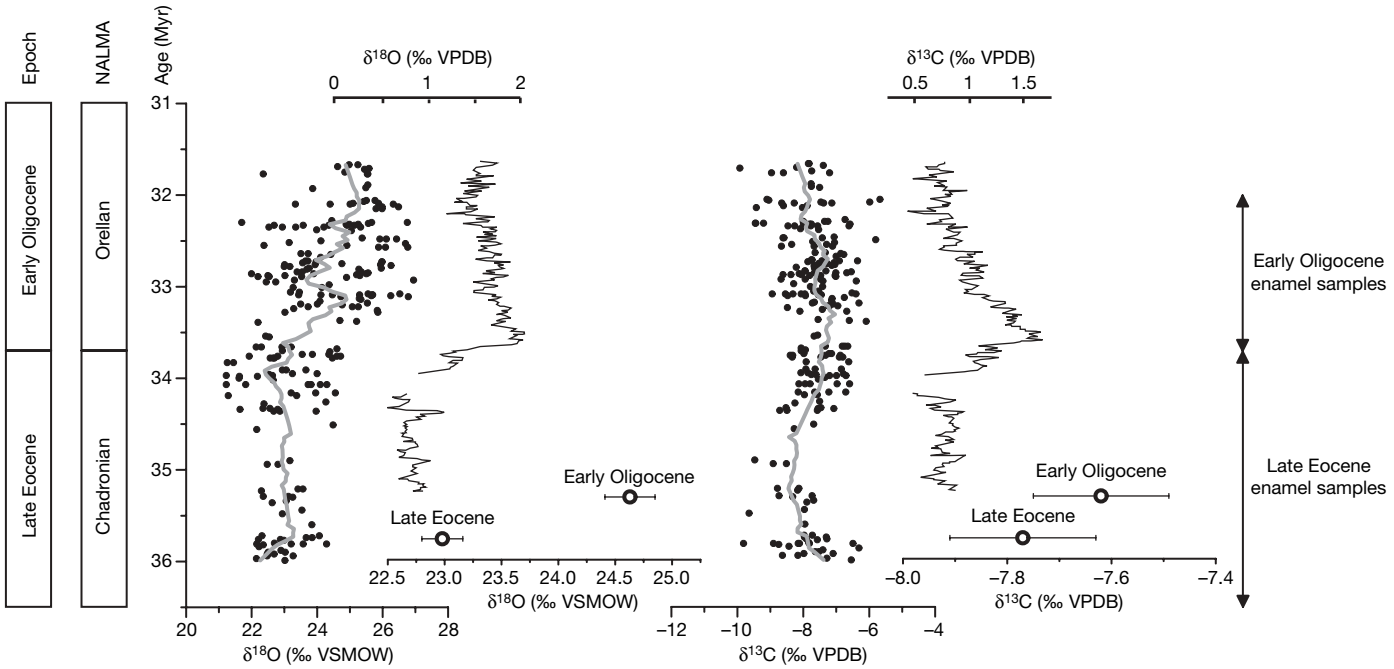


Figure 3 | High-resolution bone carbonate $\delta^{18}O$ and $\delta^{13}C$ plotted versus time for the time interval 36–32 Myr ago. The grey curves are obtained by exponential smoothing of the raw data. The open circles show the average Eocene and Oligocene $\delta^{18}O$ and $\delta^{13}C$ (on the axes directly below) with the error bars corresponding to ± 2 s.e. The new high-resolution marine curve from ref. 16 is plotted (solid line) along with our continental data (filled circles). The time interval covered by the tooth enamel samples is also shown. Bone $\delta^{13}C$ values show no consistent trend, suggesting no major change in water stress and canopy structure across the EOT. In contrast, bone $\delta^{18}O$ values show an abrupt, $\sim$1.7‰, statistically significant increase across the transition. This increase translates into a $\sim 8\ ^{\circ}C$ decrease in MAT. The continental isotopic shift may lag the marine transition by as much as 400 kyr. NALMA, North America Land-Mammal Ages.

turnovers or extinctions of gastropods, amphibians and reptiles associated with the EOT. Relative to the oceanic record, the continental transition is of higher magnitude and may lag the marine transition by as much as 400 kyr, possibly indicating a decoupling of the terrestrial and oceanic components of the climate system during major episodes of climate change.

Received 1 August; accepted 18 December 2006.

1. Miller, K. G., Fairbanks, R. G. & Mountain, G. S. Tertiary oxygen isotope synthesis, sea level history, and continental margin erosion. *Paleoceanography* **2,** 1–19 (1987).
2. Hutchison, J. H. in *Eocene-Oligocene Climatic and Biotic Evolution* (eds Prothero, D. R. & Berggren, W. A.) 451–463 (Princeton Univ. Press, Princeton, New Jersey, 1992).
3. Prothero, D. R. & Heaton, T. H. Faunal stability during the early Oligocene climatic crash. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **127,** 257–283 (1996).
4. Retallack, G. J. in *Eocene-Oligocene Climatic and Biotic Evolution* (eds Prothero, D. R. & Berggren, W. A.) 382–398 (Princeton Univ. Press, Princeton, New Jersey, 1992).
5. Wolfe, J. A. Tertiary climatic changes at middle latitudes of western North America. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **108,** 195–205 (1994).
6. Terry, D. O. Paleopedology of the Chadron Formation of northwestern Nebraska: implications for paleoclimatic change in the North American midcontinent across the Eocene-Oligocene boundary. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **168,** 1–38 (2001).
7. Bestland, E. A., Retallack, G. J. & Swisher, C. C. Stepwise climate change recorded in Eocene-Oligocene paleosol sequences from central Oregon. *J. Geol.* **105,** 153–172 (1997).
8. Evanoff, E., Prothero, D. R. & Lander, R. H. in *Eocene-Oligocene Climatic and Biotic Evolution* (eds Prothero, D. R. & Berggren, W. A.) 116–130 (Princeton Univ. Press, Princeton, New Jersey, 1992).
9. Zachos, J. C., Stott, L. D. & Lohmann, K. C. Evolution of early Cenozoic marine temperatures. *Paleoceanography* **9,** 353–387 (1994).
10. Lear, C. H., Elderfield, H. & Wilson, P. A. Cenozoic deep-sea temperatures and global ice volumes from Mg/Ca in benthic foraminiferal calcite. *Science* **287,** 269–272 (2000).
11. Kennett, J. P. Cenozoic evolution of Antarctic glaciation, the circum-Antarctic ocean, and their impact on global paleoceanography. *J. Geophys. Res.* **87,** 3843–3860 (1977).
12. DeConto, R. M. & Pollard, D. Rapid Cenozoic glaciation of Antarctica induced by declining atmospheric CO_2. *Nature* **421,** 245–249 (2003).
13. Aubry, M. P. in *Eocene-Oligocene Climatic and Biotic Evolution* (eds Prothero, D. R. & Berggren, W. A.) 272–309 (Princeton Univ. Press, Princeton, New Jersey, 1992).
14. Baldauf, J. G. in *Eocene-Oligocene Climatic and Biotic Evolution* (eds Prothero, D. R. & Berggren, W. A.) 310–326 (Princeton Univ. Press, Princeton, New Jersey, 1992).
15. Hansen, T. A., Kelley, P. H. & Haasl, D. M. Paleoecological patterns in molluscan extinctions and recoveries: comparison of the Cretaceous-Paleogene and Eocene-Oligocene extinctions in North America. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **214,** 233–242 (2004).
16. Coxall, H. K., Wilson, P. A., Pälike, H., Lear, C. H. & Backman, J. Rapid stepwise onset of Antarctic glaciation and deeper calcite compensation in the Pacific Ocean. *Nature* **433,** 53–57 (2005).
17. Swisher, C. C. & Prothero, D. R. Single crystal $^{40}Ar/^{39}Ar$ dating of the Eocene-Oligocene transition in North America. *Science* **249,** 760–762 (1990).
18. Prothero, D. R. & Whittlesey, K. E. in *Depositional Environments, Lithostratigraphy, and Biostratigraphy of the White River and Arikaree Groups (Late Eocene to early Miocene, North America)* (eds Terry, D. O., LaGarry, H. E. & Hunt, R. M.) 39–62 (Geol. Soc. Am. Spec. Pap. 325, The Geological Society of America, Boulder, Colorado, 1998).
19. Prothero, D. R. in *The Terrestrial Eocene-Oligocene Transition in North America* (eds Prothero, D. R. & Emry, R. J.) 262–277 (Cambridge Univ. Press, Cambridge, UK, 1996).
20. Kohn, M. J. & Cerling, T. E. Stable isotope compositions of biological apatite. *Rev. Mineral.* **48,** 455–488 (2002).
21. Kohn, M. J. & Law, J. M. Stable isotope chemistry of fossil bone as a new paleoclimate indicator. *Geochim. Cosmochim. Acta* **70,** 931–946 (2006).
22. Cerling, T. E. & Quade, J. in *Climate Change in Continental Isotopic Records* (eds Swart, P. K., Lohmann, K. C., McKenzie, J. & Savin, S.) 217–231 (Geophys. Monogr. 78, AGU, Washington DC, 1993).
23. Cerling, T. E. *et al.* Global vegetation change through the Miocene/Pliocene boundary. *Nature* **389,** 153–158 (1997).
24. Kohn, M. J. & Fremd, T. J. Tectonic controls on isotope compositions and species diversification, John Day Basin, central Oregon. *PaleoBios* (in the press).
25. Janis, C. The evolutionary strategy of the equidae and the origins of rumen and cecal digestion. *Evolution Int. J. Org. Evolution* **30,** 757–774 (1976).
26. Chase, C. G., Gregory-Wodzicki, K. M., Parrish, J. T. & DeCelles, P. G. in *Tectonic Boundary Conditions for Climate Model Simulations* (eds Crowley, T.J. & Burke, K.) 73–99 (Oxford Univ. Press, Oxford, UK, 1998).
27. Ivany, L. C., Patterson, W. P. & Lohmann, K. C. Cooler winters as a possible cause of mass extinctions at the Eocene/Oligocene boundary. *Nature* **407,** 887–890 (2000).
28. Hendricks, M. B., DePaolo, D. J. & Cohen, R. C. Space and time variation of $\delta^{18}O$ and δD in precipitation: can paleotemperature be estimated from ice cores? *Glob. Biogeochem. Cycles* **14,** 851–861 (2000).

29. Toggweiler, J. R. & Bjornsson, H. Drake passage and paleoclimate. *J. Quat. Sci.* **15,** 319–328 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the NSF. Fossil tooth samples were provided by the University of Florida (UF), the Denver Museum of Nature and Science, and the American Museum of Natural History. Fossils sampled from the UF were collected on public lands under the auspices of a US Forest Service permit or on private lands with the kind permission of the leaseholders B. and R. Toomey. We thank E. Tappa and W. Straight for facilitating analyses, D. Fox and J. Mintz for their help in the field, and B. Beasley, M. Jin, K. Carpenter and R. Graham for allowing the sampling of bones and teeth.

Author Contributions A.Z. and M.J.K. conceived the project, collected samples and data, and led write-up. B.J.M. collected tooth samples at Toadstool Park. D.O.T. helped collect bone samples and refine chronology and stratigraphy at Toadstool Park. All authors discussed and commented on this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.Z. (azanazzi@geol.sc.edu) or M.J.K. (mjk@geol.sc.edu).

LETTERS

Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models

Anatol C. Kreitzer[1]† & Robert C. Malenka[1]

The striatum is a major forebrain nucleus that integrates cortical and thalamic afferents and forms the input nucleus of the basal ganglia[1,2]. Striatal projection neurons target the substantia nigra pars reticulata (direct pathway) or the lateral globus pallidus (indirect pathway). Imbalances between neural activity in these two pathways have been proposed to underlie the profound motor deficits observed in Parkinson's disease and Huntington's disease[3,4]. However, little is known about differences in cellular and synaptic properties in these circuits. Indeed, current hypotheses suggest that these cells express similar forms of synaptic plasticity[5,6]. Here we show that excitatory synapses onto indirect-pathway medium spiny neurons (MSNs) exhibit higher release probability and larger *N*-methyl-D-aspartate receptor currents than direct-pathway synapses. Moreover, indirect-pathway MSNs selectively express endocannabinoid-mediated long-term depression (eCB-LTD), which requires dopamine D2 receptor activation. In models of Parkinson's disease, indirect-pathway eCB-LTD is absent but is rescued by a D2 receptor agonist or inhibitors of endocannabinoid degradation. Administration of these drugs together *in vivo* reduces parkinsonian motor deficits, suggesting that endocannabinoid-mediated depression of indirect-pathway synapses has a critical role in the control of movement. These findings have implications for understanding the normal functions of the basal ganglia, and also suggest approaches for the development of therapeutic drugs for the treatment of striatal-based brain disorders.

Experimental and clinical work based on anatomical and molecular differences in indirect-pathway and direct-pathway MSNs has led to the suggestion that these pathways contribute differentially to the pathophysiology of striatal disorders[3,4,7–9]. However, progress has been hampered by a lack of knowledge about possible physiological differences between direct-pathway and indirect-pathway MSNs. To compare the cellular and synaptic properties of these two classes of MSN we took advantage of bacterial artificial chromosome (BAC) transgenic mice that confer cell-type-specific expression of green fluorescent protein (GFP) in distinct neuronal subpopulations, including MSNs of the direct and indirect pathways[6,10,11]. Expression of GFP from the muscarinic M4 receptor locus labelled striatonigral MSNs of the direct pathway (M4–GFP; Fig. 1b). In contrast, GFP expression driven from the dopamine D2 receptor locus labelled striatopallidal MSNs of the indirect pathway (D2–GFP, Fig. 1c).

Whole-cell voltage-clamp recordings from GFP-positive and GFP-negative MSNs in slices from the M4–GFP line revealed that synapses onto direct-pathway MSNs had larger paired-pulse ratios (PPRs) than synapses onto indirect-pathway MSNs (direct-pathway PPRs at 50 ms, 1.28 ± 0.06, $n = 11$; indirect-pathway PPRs at 50 ms, 1.1 ± 0.04, $n = 7$; $P < 0.05$; Fig. 1d, e). This suggests that indirect-pathway synapses have a higher probability of neurotransmitter release than direct-pathway synapses[12]. To confirm that the M4–GFP and D2–GFP mouse lines label complementary populations of MSNs, we conducted similar experiments in the D2–GFP mouse line. Consistent with our results was our observation that excitatory synapses on direct-pathway MSNs (now GFP negative) in the D2–GFP mouse line had larger PPRs than indirect-pathway (now GFP positive) synapses (direct-pathway PPRs at 50 ms, 1.41 ± 0.06, $n = 10$; indirect-pathway PPRs at 50 ms, 1.1 ± 0.07, $n = 9$; $P < 0.05$; Fig. 1f, g). To examine synaptic properties further in these two cell populations (Supplementary Methods), we recorded miniature excitatory postsynaptic currents (mEPSCs) (Fig. 1h). Direct-pathway neurons had a lower mEPSC frequency than indirect-pathway neurons (direct, 1.7 ± 0.1 Hz, $n = 4$; indirect, 3.7 ± 0.7 Hz, $n = 5$; $P < 0.05$; Fig. 1i), whereas no differences were observed in mEPSC amplitudes across the two populations (direct, 17.1 ± 1 pA, $n = 4$; indirect, 15.1 ± 0.9 pA, $n = 5$; $P > 0.05$; Fig. 1j). These results provide further evidence that synapses on indirect-pathway MSNs have a higher probability of neurotransmitter release than synapses on direct-pathway MSNs.

The lack of differences in mEPSC amplitude suggests that the α-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA) receptor densities at synapses on the two populations of MSNs are similar. To determine whether the subunit composition of synaptic AMPA receptors might differ[13], we recorded EPSCs at a range of membrane potentials. In both direct-pathway and indirect-pathway MSNs, EPSCs showed moderate inward rectification (current ratio between $+50$ mV and -50 mV: direct, 0.51 ± 0.07, $n = 4$; indirect, 0.48 ± 0.03, $n = 5$; Fig. 2a) as well as similar decay time constants (direct, 5.2 ± 0.1 ms, $n = 18$; indirect, 5.4 ± 0.1 ms, $n = 15$), probably indicating a mixture of GluR2-containing and GluR2-lacking AMPA receptors in both MSN populations. To examine possible differences in synaptic *N*-methyl-D-aspartate (NMDA) receptors, we measured the ratio of NMDA receptor-mediated synaptic currents (current at 50 ms, $V_{hold} = +40$ mV) to AMPA receptor-mediated synaptic currents (peak current, $V_{hold} = -60$ mV). Indirect-pathway synapses had significantly larger NMDA/AMPA ratios than direct-pathway synapses (direct, 0.39 ± 0.05, $n = 14$; indirect, 0.53 ± 0.04, $n = 14$; $P < 0.05$; Fig. 2b–d), although the decay time constants of NMDA receptor-mediated EPSCs were similar in both populations (direct, 51.3 ± 0.5 ms, $n = 15$; indirect, 49.5 ± 0.7 ms, $n = 14$). These findings suggest that indirect-pathway synapses have a higher density of NMDA receptors than direct-pathway synapses but the subunit compositions of synaptic NMDA receptors in the two populations are similar.

We next examined whether the intrinsic membrane excitability of direct-pathway and indirect-pathway MSNs differed. Although the resting membrane potentials and input resistances of the two populations were similar (direct, -91 ± 1 mV, 51 ± 8 MΩ, $n = 7$; indirect, -91 ± 1 mV, 64 ± 8 MΩ, $n = 8$; $P > 0.05$ for both measures) indirect-pathway MSNs fired at nearly twice the rate of direct-pathway neurons in response to depolarizing current injection (Fig. 2e–g). No

[1]Department of Psychiatry and Behavioral Sciences, Nancy Pritzker Laboratory, Stanford University Medical School, Palo Alto, California 94305, USA. †Present address: Gladstone Institute of Neurological Disease and Department of Physiology, University of California, San Francisco, San Francisco, California 94158, USA.

adaptation in the frequency of spiking during the depolarization was observed in either population (Fig. 2h).

These electrophysiological differences between indirect-pathway and direct-pathway MSNs may be important for spike integration, downstate to upstate transitions, and the induction of synaptic plasticity. For example, the higher probability of release at indirect-pathway synapses might enhance the activation of metabotropic glutamate receptors (mGluRs), whereas the increased excitability of indirect-pathway MSNs may increase the activation of L-type calcium channels. Indeed, the most prominent form of synaptic plasticity in the striatum is LTD[5,14,15], which requires the activation of L-type calcium channels and mGluRs, leading to the release of endocannabinoids and a long-lasting inhibition of neurotransmitter release. That striatal LTD also requires D2 receptor activation provides further support for the idea that it may be more robust at indirect-pathway MSNs[16–18]. We therefore examined the properties of LTD at direct-pathway and indirect-pathway synapses.

High-frequency afferent stimulation caused robust LTD at indirect-pathway MSN synapses (51 ± 5% of baseline at 30–40 min, $n = 6$, $P < 0.05$; Fig. 3a, b). However, no LTD was apparent at direct-pathway synapses (101 ± 10% of baseline at 30–40 min, $n = 12$, $P > 0.05$; Fig. 3a, b) in response to the same induction protocol. The LTD in indirect-pathway MSNs was blocked by the cannabinoid CB1 receptor antagonist AM251 (1 μM) (107 ± 10% of baseline at 30–40 min, $n = 5$, $P > 0.05$; Fig. 3c), confirming that it was mediated by endocannabinoids. It was also blocked by the D2 receptor antagonist sulpiride (10 μM) (96 ± 12% of baseline at 30–40 min, $n = 6$, $P > 0.05$; Fig. 3c) but not by a battery of serotonin and noradrenaline receptor antagonists (methysergide (4 μM), GR125487 (1 μM), LY278584 (1 μM), prazosin (2 μM) and propranolol (1 μM)) (50 ± 13% of baseline at 30–40 min, $n = 5$, $P < 0.05$; Fig. 3c). These results provide further evidence that endogenous dopamine, but not other monoamines, is critical for the triggering of eCB-LTD. The magnitude of eCB-LTD in indirect-pathway MSNs is nearly twice as large as reported previously[16], suggesting that in previous studies data from indirect-pathway and direct-pathway MSNs were averaged.

The lack of eCB-LTD in direct-pathway MSN synapses could be due to a lack of presynaptic CB1 receptors on the corresponding presynaptic terminals. However, application of the CB1 receptor agonist WIN55,212 (1 μM) decreased EPSCs at both direct-pathway and indirect-pathway synapses to a similar extent (direct, 56 ± 8% of baseline at 15–20 min after wash-in, $n = 4$, $P < 0.05$; indirect, 47 ± 7% of baseline at 15–20 min after wash-in, $n = 4$, $P < 0.05$; Fig. 3d). This finding suggests that the postsynaptic biosynthesis and/or release of endogenous cannabinoids is different between indirect-pathway and direct-pathway MSNs. To test this hypothesis, we applied the type I mGluR agonist (*S*)-3,5-dihydroxyphenylglycine (DHPG; 100 μM for 10 min) while holding MSNs at a slightly depolarized membrane potential (−50 mV), a manipulation that elicits endocannabinoid release from MSNs and triggers eCB-LTD[16]. This protocol yielded robust LTD at indirect-pathway synapses (64 ± 5% of baseline at 30–40 min, $n = 5$, $P < 0.05$; Fig. 3e) but only a small and reversible depression at direct-pathway synapses (95 ± 9% of

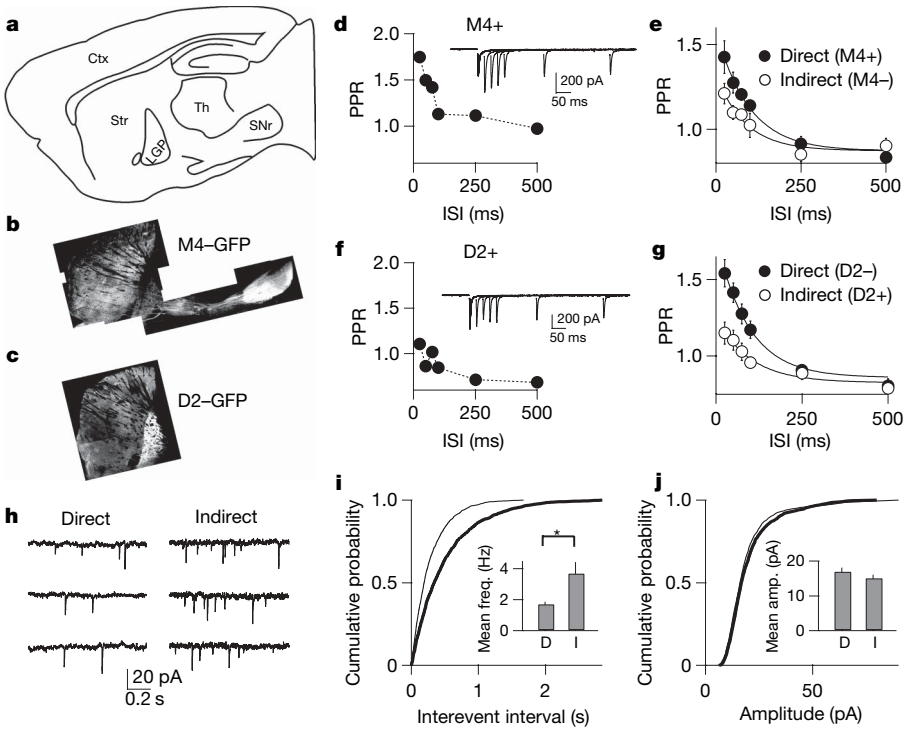


Figure 1 | Presynaptic properties of direct-pathway and indirect-pathway synapses on striatal MSNs. **a**, Sagittal diagram of mouse brain showing cortex (Ctx), striatum (Str), lateral globus pallidus (LGP), thalamus (Th) and substantia nigra pars reticulata (SNr). **b**, Composite confocal image of GFP fluorescence in a sagittal section from an M4–GFP BAC-transgenic mouse. **c**, Composite confocal image of GFP fluorescence in a sagittal section from a D2–GFP BAC-transgenic mouse. **d**, Representative recording from a GFP-positive MSN in an M4–GFP mouse. PPRs ($EPSC_2/EPSC_1$) are plotted against interstimulus interval (ISI). Points represent averages of three or four trials (filled circles). Inset: traces from an individual trial. **e**, Summary graph of PPRs from GFP-positive (direct pathway, filled circles; $n = 11$) and GFP-negative (indirect pathway, open circles; $n = 7$) neurons in M4–GFP mice plotted against ISI. **f**, Representative recording from a GFP-positive MSN in a D2–GFP mouse. PPRs are plotted against ISI. Points represent averages of three or four trials (filled circles). Inset: traces from an individual trial. **g**, Summary graph of PPRs from GFP-positive (indirect, open circles; $n = 9$) and GFP-negative (direct pathway, filled circles; $n = 10$) neurons in D2–GFP mice, plotted against ISI. **h**, Sample traces of mEPSCs recorded in tetrodotoxin (1 μM) from direct-pathway and indirect-pathway neurons. **i**, Cumulative probability plots for mEPSC inter-event interval in direct-pathway (thick trace; $n = 4$) and indirect-pathway (thin trace; $n = 5$) neurons. Inset: mean frequency for direct-pathway (D) and indirect-pathway (I) neurons. **j**, Cumulative probability plot for mEPSC amplitude in direct-pathway (thick trace) and indirect-pathway (thin trace) neurons. Inset: mean amplitude for direct (D) and indirect (I) pathways. Asterisk, $P < 0.05$. Error bars indicate s.e.m.

baseline at 30–40 min, $n = 6$, $P > 0.05$). Because dopamine acting through D2 receptors enhances striatal endocannabinoid release[19] and D2 receptor activation is necessary for striatal LTD[5,16,18], we next asked whether the D2 receptor agonist quinpirole might enhance the depression of EPSCs in indirect-pathway MSNs in response to a brief (1-min) application of DHPG, neither of which alone is sufficient to elicit LTD[16]. As expected, brief DHPG application alone did not yield a lasting depression of EPSCs (98 ± 5% of baseline at 20–30 min, $n = 5$, $P > 0.05$), but in the presence of quinpirole (10 μM) the same DHPG application generated LTD (73 ± 8% of baseline at 20–30 min, $n = 6$, $P < 0.05$; Fig. 3f).

These results show that direct-pathway and indirect-pathway MSNs respond differently to both synaptic stimulation and mGluR activation, suggesting that endocannabinoid release sufficient to trigger eCB-LTD is restricted to indirect-pathway MSNs. The ability to observe LTD at most MSNs in previous studies[6,14] may have been due to differences in experimental protocols that led to a spill-over of endocannabinoids from indirect-pathway postsynaptic neurons onto direct-pathway presynaptic terminals. However, we have not observed such spill-over under a variety of stimulation and recording conditions, nor did we observe a depression of corticostriatal transmission in response to the muscarinic receptor antagonist pirenzepine[6] (Supplementary Fig. S1).

To investigate whether dopamine depletion would block eCB-LTD at indirect-pathway MSNs, we used two animal models of Parkinson's disease, treatment with reserpine and with 6-hydroxydopamine (6-OHDA)[20,21]. LTD was not elicited in indirect-pathway MSNs in slices prepared from mice that had received reserpine (5 mg kg^{-1} intraperitoneally (i.p.)) 18–24 h earlier (122 ± 8% of baseline at 30–40 min, $n = 6$; Fig. 4a), a result providing further evidence that dopamine is critical for triggering eCB-LTD. (Reserpine also decreases the levels of other monoamines, but the data in Fig. 3c show that these monoamines are not required for eCB-LTD.) To determine whether eCB-LTD could be rescued in reserpine-treated mice, we applied tetanic stimulation in the presence of quinpirole (10 μM) and found that this facilitated the generation of eCB-LTD at indirect-pathway synapses (81 ± 9% of baseline at 30–40 min, $n = 5$, $P < 0.05$ reserpine-treated versus reserpine-treated with quinpirole; Fig. 4a). We could also rescue LTD by the application of URB597 (1 μM), a potent inhibitor of fatty acid amide hydrolase (FAAH), the degradative enzyme for the endogenous cannabinoid anandamide[22] (62 ± 5% of baseline at 30–40 min, $n = 5$, $P < 0.05$ reserpine-treated versus reserpine-treated with URB597; Fig. 4c). Application of URB597 alone had no effect on EPSC amplitudes ($n = 4$; data not shown).

We obtained similar results with mice that had received bilateral 6-OHDA injections into the medial forebrain bundle 48–60 h before slice preparation. Specifically, tetanic stimulation did not elicit LTD

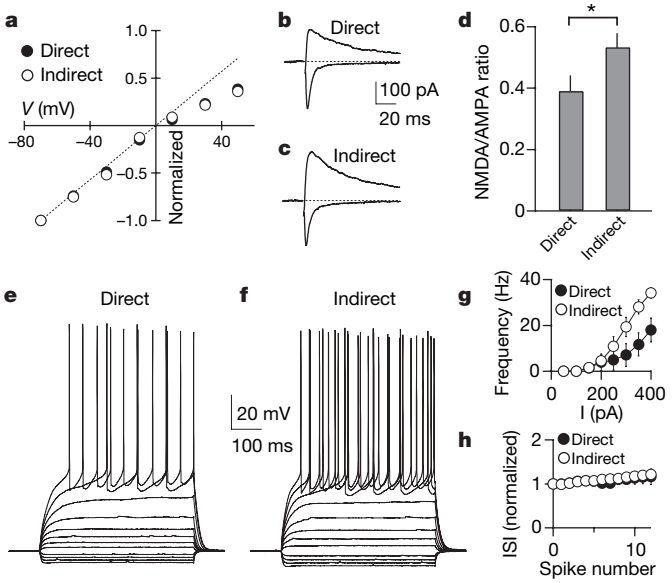


Figure 2 | Postsynaptic properties of direct-pathway and indirect-pathway synapses. a, Summary of normalized EPSC amplitudes recorded in APV (50 μM) with spermine (0.1 mM) in the pipette at different holding potentials in direct-pathway (filled circles; $n = 4$) and indirect-pathway (open circles; $n = 5$) MSNs. **b**, EPSC traces recorded at –60 mV and +40 mV from a representative direct-pathway MSN. **c**, EPSC traces recorded at –60 mV and +40 mV from a representative indirect-pathway MSN. **d**, Summary of the ratio of the NMDA-receptor EPSC (measured at 50 ms after stimulus, holding potential: +40 mV) to the AMPA-receptor EPSC (measured at peak, holding potential −60 mV) at direct-pathway ($n = 14$) and indirect-pathway ($n = 14$) synapses. **e, f**, Current-clamp recording from a representative direct-pathway (**e**) or indirect-pathway (**f**) MSN displaying responses to injected current (50-pA steps). **g**, Summary of firing frequency in response to injected current in direct-pathway (filled circles; $n = 7$) and indirect-pathway (open circles; $n = 8$) neurons. **h**, Summary of normalized interspike interval (ISI) plotted as a function of spike number. Filled circles, direct; open circles, indirect. Asterisk, $P < 0.05$. Error bars indicate s.e.m.

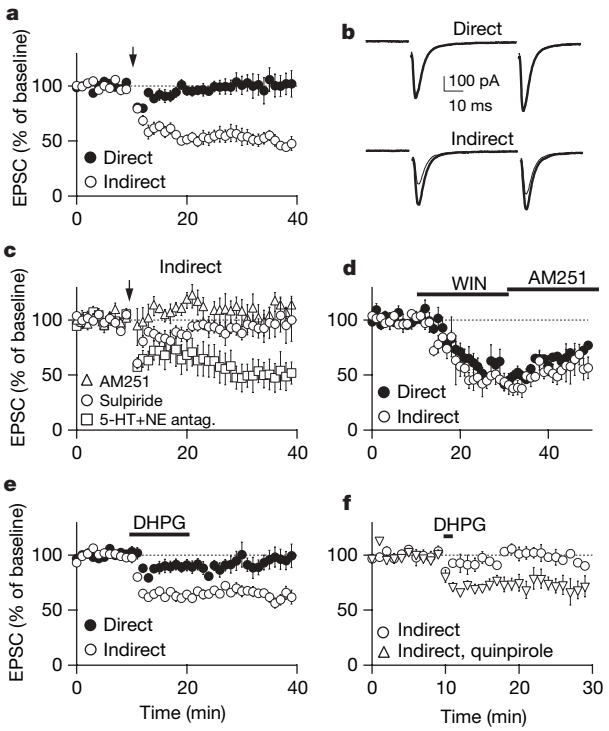


Figure 3 | Endocannabinoid-mediated LTD is restricted to indirect-pathway synapses. a, LTD is present in indirect-pathway (open circles; $n = 6$) but not direct-pathway (filled circles; $n = 12$) MSNs. In this and subsequent panels, normalized EPSCs recorded from direct-pathway and indirect-pathway MSNs are plotted over time. The arrow indicates 1 s of 100 Hz stimulation, paired with postsynaptic depolarization to 0 mV, repeated four times at 10-s intervals. **b**, Average traces from a representative LTD experiment in a direct (top) and indirect (bottom) pathway MSN. Thick traces represent the average from 0–10 min, and thin traces represent the average from 30–40 min. **c**, LTD in indirect-pathway neurons is blocked by sulpiride (10 μM) (circles; $n = 6$) and AM251 (1 μM) (triangles; $n = 5$) but not by a combination of methysergide (4 μM), GR125487 (1 μM), LY278584 (1 μM), prazosin (2 μM) and propranolol (1 μM) (squares; $n = 5$). **d**, Activation of CB1 receptors inhibits neurotransmitter release at both direct-pathway (filled circles; $n = 4$) and indirect-pathway (open circles; $n = 4$) synapses. WIN55,212-2 (1 μM) and AM251 (5 μM) were applied during the time marked by the black bars. **e**, Application of DHPG (100 μM) for 10 min elicits LTD in indirect-pathway (open circles; $n = 5$) but not direct-pathway (filled circles; $n = 6$) MSNs. **f**, Application of DHPG (100 μM) for 1 min does not elicit LTD at indirect-pathway synapses (circles; $n = 5$), but does elicit LTD in the presence of quinpirole (10 μM) (triangles; $n = 6$). Error bars indicate s.e.m.

at indirect-pathway synapses in slices prepared from 6-OHDA-treated mice (111 ± 13% of baseline at 30–40 min, $n = 7$, $P > 0.05$; Fig. 4b) but was rescued by applying either quinpirole (10 μM) (65 ± 8% of baseline at 30–40 min, $n = 4$, $P < 0.05$ 6-OHDA-treated versus 6-OHDA-treated with quinpirole; Fig. 4b) or URB597 (1 μM) (55 ± 7% of baseline at 30–40 min, $n = 4$, $P < 0.05$ 6-OHDA-treated versus 6-OHDA-treated with URB597; Fig. 4d). Thus, D2 receptor activation or inhibition of endocannabinoid degradation can rescue indirect-pathway eCB-LTD in dopamine-depleted animals. Although biochemical measurements have suggested that endocannabinoid levels in striatal brain tissue may increase after dopamine depletion[23,24], we find no physiological evidence consistent with increased endocannabinoid levels in dopamine-depleted animals (Supplementary Fig. S2).

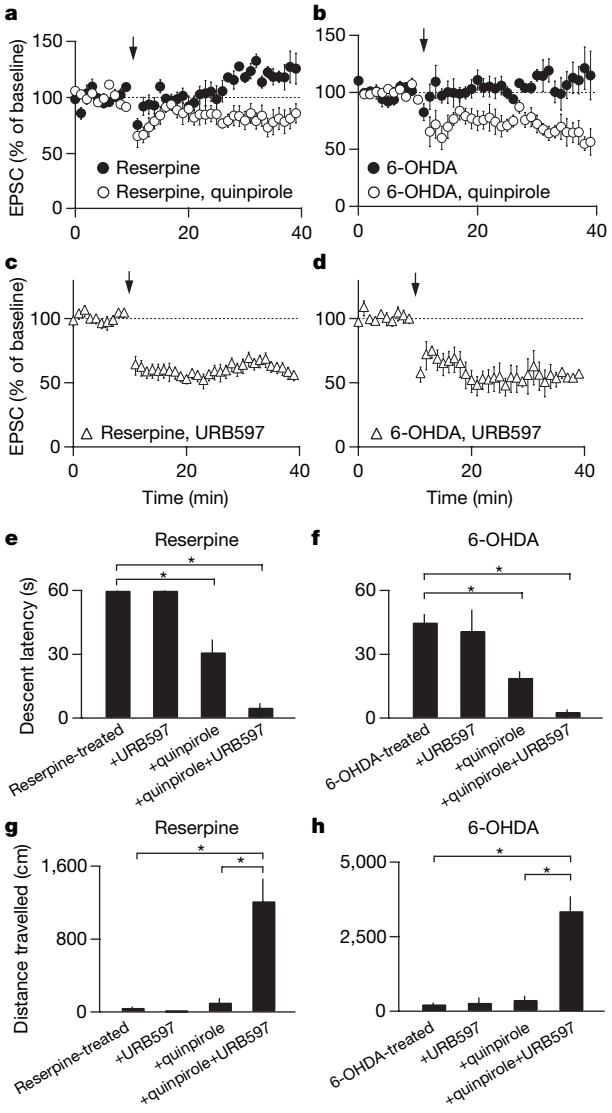


Figure 4 | Pharmacological rescue of indirect-pathway LTD and motor deficits in animal models of Parkinson's disease. **a**, **b**, LTD is absent (filled circles) in reserpine-treated mice ($n = 6$) (**a**) or 6-OHDA-treated mice ($n = 7$) (**b**), but is partly rescued when quinpirole (10 μM) (open circles) is present during the induction protocol in reserpine-treated mice ($n = 5$) (**a**) or 6-OHDA-treated mice ($n = 4$) (**b**). **c**, **d**, LTD is rescued in reserpine-treated mice ($n = 5$) (**c**) or 6-OHDA-treated mice ($n = 4$) (**d**) in the presence of URB597 (1 μM). Normalized EPSCs recorded from indirect-pathway MSNs are plotted over time. The arrow above the graphs in **a**–**d** indicates the time of the LTD induction protocol. **e**, **f**, Descent latency in the catalepsy bar test for reserpine-treated mice (**e**) and 6-OHDA-treated mice (**f**) before subsequent drug treatment (reserpine, $n = 14$; 6-OHDA, $n = 20$), and after injection with URB597 (1 mg kg^{-1} i.p.; reserpine, $n = 4$; 6-OHDA, $n = 5$), quinpirole (1.5 mg kg^{-1} i.p.; reserpine, $n = 5$; 6-OHDA, $n = 10$), or URB597 and quinpirole together (reserpine, $n = 7$; 6-OHDA, $n = 5$). **g**, **h**, Locomotor activity in the open-field test for animals treated with reserpine (**g**) or 6-OHDA (**h**) plotted as distance travelled during a 15-min test period before subsequent drug treatment (reserpine, $n = 19$; 6-OHDA, $n = 20$), and for mice 45–60 min after injection with URB597 (1 mg kg^{-1} i.p.; reserpine, $n = 4$; 6-OHDA, $n = 5$), quinpirole (1.5 mg kg^{-1} i.p.; reserpine, $n = 8$; 6-OHDA, $n = 5$), or URB597 and quinpirole together (reserpine, $n = 7$; 6-OHDA, $n = 5$). Asterisk, $P < 0.05$ by one-way analysis of variance with Tukey's Honestly Significant Difference or Dunnett's test. Error bars indicate s.e.m.

The success of therapeutic interventions that activate D2 receptors and reduce indirect-pathway activity in Parkinson's disease[8,25] provided the motivation to test whether inhibiting endocannabinoid degradation, either alone or in combination with D2 receptor activation, could improve the motor deficits in mice treated with reserpine or with 6-OHDA. At 18–24 h after reserpine injections (1 mg kg^{-1} i.p.), mice displayed pronounced catalepsy (descent latency: more than 60 s, $n = 14$; Fig. 4e) and locomotor activity was minimal (distance travelled in 15 min: 48 ± 11 cm, $n = 19$; Fig. 4g). Similarly, at 48–60 h after bilateral injection of 6-OHDA into the medial forebrain bundle, mice exhibited catalepsy (descent latency: 45 ± 4 s, $n = 20$; Fig. 4f) and minimal open field locomotor activity (distance travelled: 236 ± 52 cm, $n = 15$; Fig. 4h). After administration of URB597 alone (1 mg kg^{-1} i.p.), no significant decrease in catalepsy or increase in locomotor activity was observed in mice treated with either reserpine or 6-OHDA. Injection of the D2 agonist quinpirole alone (1.5 mg kg^{-1} i.p.) decreased catalepsy (reserpine-treated mice, descent latency: 31 ± 6 s, $n = 5$, $P < 0.05$; 6-OHDA-treated mice, descent latency: 19 ± 3 s, $n = 10$, $P < 0.05$) but had no significant effect on locomotor activity (reserpine-treated mice, distance travelled: 105 ± 47 cm, $n = 8$, $P > 0.05$; 6-OHDA-treated mice, distance travelled: 385 ± 129 cm, $n = 5$; $P > 0.05$). However, when URB597 was administered together with quinpirole, catalepsy was markedly decreased (reserpine-treated mice, descent latency: 5 ± 2 s, $n = 7$, $P < 0.05$; 6-OHDA-treated mice, descent latency: 3 ± 1 s, $n = 5$, $P < 0.05$) and locomotor activity was increased significantly more than with quinpirole alone (reserpine-treated mice, distance travelled: 1,217 ± 248 cm, $n = 7$, $P < 0.05$; 6-OHDA-treated mice, distance travelled: 3,363 ± 490 cm, $n = 5$, $P < 0.05$). Thus, the effects of these drugs on indirect-pathway eCB-LTD *in vitro* were predictive of their therapeutic efficacy in two different models of Parkinson's disease.

In vivo in the absence of dopamine, endogenous excitatory activity may not generate enough endocannabinoid release from indirect-pathway MSNs to yield significant LTD even in the presence of inhibitors of endocannabinoid degradation. However, when endocannabinoids are more effectively released as a result of D2 receptor activation by quinpirole, we propose that the endocannabinoid degradation inhibitor URB597 enhances LTD induction, thereby reducing indirect-pathway neuron activity and more effectively restoring movement. Similar electrophysiological and behavioural results were obtained with URB754 (ref. 26) (Supplementary Figs S3 and S4), a putative inhibitor of the endocannabinoid-degrading enzyme monoacylglycerol lipase, although the specificity of this drug has recently been called into question[27].

We have found major differences in the cellular and synaptic properties of striatal MSNs in the direct and indirect basal ganglia pathways. Most notably, indirect-pathway MSNs are more excitable and selectively express dopamine-dependent and endocannabinoid-dependent LTD (Supplementary Fig. S5). Dopamine depletion in animal models of Parkinson's disease blocked the generation of eCB-LTD, whereas its rescue by a dopamine D2 receptor agonist or an inhibitor of endocannabinoid degradation predicted the therapeutic benefits of these agents in improving parkinsonian motor deficits *in vivo*. Consistent with these findings are the observations that both D2 receptor and CB1 receptor knockout mice lack striatal LTD[15,17,18] and exhibit profound motor deficits similar to those

observed in Parkinson's disease[28,29]. However, the administration of CB1 receptor antagonists alone does not reduce locomotor activity[30], and the depletion of dopamine during Parkinson's disease probably results in a host of changes in the cellular and synaptic properties of both direct-pathway and indirect-pathway MSNs. Thus, the elimination of eCB-LTD at indirect-pathway synapses may be only one component of the complex pathophysiology of Parkinson's disease.

The ability to generate eCB-LTD in indirect-pathway MSNs but not direct-pathway MSNs seems to be due to differences in the postsynaptic generation of endocannabinoids in response to synaptic activity. Indeed, recent gene profiling of direct-pathway and indirect-pathway MSNs indicates that there are marked differences in the complement of signal transduction molecules and G-protein-coupled receptors that they express[11]. The identification of such cell-type-restricted proteins in the striatum along with the characterization of their physiological functions should facilitate the discovery of new drug targets with the therapeutic advantage of allowing specific and independent control of indirect-pathway and direct-pathway activity. Together with previous results, our findings specifically suggest that manipulation of activity in the indirect basal ganglia pathway by means of modulation of endocannabinoid production may be particularly beneficial for brain disorders that involve dysfunctions of striatal circuitry, such as Parkinson's disease.

METHODS

See Supplementary Information for detailed methods.

Electrophysiology. Coronal brain slices (300 μm) were prepared from M4–GFP or D2–GFP heterozygotic BAC-transgenic mice (postnatal days 20–25). Whole-cell voltage-clamp and current-clamp recordings were obtained from visually identified GFP-positive or GFP-negative MSNs in dorsolateral striatum at a temperature of 30–32 °C, with picrotoxin (50 μM) present to suppress $GABA_A$-mediated currents. Excitatory synaptic currents were evoked by intrastriatal microstimulation with a saline-filled glass pipette placed 50–100 μm dorsolateral of the recorded neuron. All data acquisition and analysis was performed online with custom Igor Pro software.

Dopamine depletion and behaviour. Reserpine was administered i.p. at a concentration of 1 or 5 mg kg^{-1}, and electrophysiological and behavioural experiments were performed 18–24 h after injections. 6-Hydroxydopamine was injected bilaterally into the median forebrain bundle at a concentration of 5 μg μl^{-1} to a total volume of 2 μL per side, and experiments were performed 48–60 h after injections. Catalepsy was measured with the bar test. Descent latency represents the time in seconds required for the mouse either to remove both paws from a bar (diameter 2 mm) placed 4 cm above the ground, or for one paw to touch the ground. Horizontal locomotor activity was measured in 15-min blocks with the ENV-510 test chamber and Activity Monitor software. Animals were tested 45–60 min after drug administration. All procedures involving animals were approved by the Institutional Animal Care and Use Committee.

Statistics. Summary data are reported as means ± s.e.m. Statistical significance was evaluated with a one-way analysis of variance, with posthoc tests for between-group significance (Tukey's Honestly Significant Difference and Dunnett's test) or the two-tailed unpaired *t*-test.

Received 25 August; accepted 5 December 2006.

1. Wilson, C. J. in *The Synaptic Organization of the Brain* (ed. Shepherd, G. M.) 329–375 (Oxford University Press, New York, 1998).
2. Bolam, J. P., Hanley, J. J., Booth, P. A. & Bevan, M. D. Synaptic organisation of the basal ganglia. *J. Anat.* **196,** 527–542 (2000).
3. Albin, R. L., Young, A. B. & Penney, J. B. The functional anatomy of basal ganglia disorders. *Trends Neurosci.* **12,** 366–375 (1989).
4. Graybiel, A. M. The basal ganglia. *Curr. Biol.* **10,** R509–R511 (2000).
5. Gerdeman, G. L., Partridge, J. G., Lupica, C. R. & Lovinger, D. M. It could be habit forming: drugs of abuse and striatal synaptic plasticity. *Trends Neurosci.* **26,** 184–192 (2003).
6. Wang, Z. *et al.* Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. *Neuron* **50,** 443–452 (2006).
7. Gerfen, C. R. *et al.* D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science* **250,** 1429–1432 (1990).
8. Obeso, J. A. *et al.* Pathophysiologic basis of surgery for Parkinson's disease. *Neurology* **55,** S7–S12 (2000).
9. Mallet, N., Ballion, B., Le Moine, C. & Gonon, F. Cortical inputs and GABA interneurons imbalance projection neurons in the striatum of parkinsonian rats. *J. Neurosci.* **26,** 3875–3884 (2006).
10. Gong, S. *et al.* A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425,** 917–925 (2003).
11. Lobo, M., Karsten, S., Gray, M., Geschwind, D. & Yang, X. FACS-array profiling of striatal projection neuron subtypes in juvenile and adult mouse brains. *Nature Neurosci.* **9,** 443–452 (2006).
12. Zucker, R. S. & Regehr, W. G. Short-term synaptic plasticity. *Annu. Rev. Physiol.* **64,** 355–405 (2002).
13. Dingledine, R., Borges, K., Bowie, D. & Traynelis, S. F. The glutamate receptor ion channels. *Pharmacol. Rev.* **51,** 7–61 (1999).
14. Calabresi, P., Maj, R., Pisani, A., Mercuri, N. B. & Bernardi, G. Long-term synaptic depression in the striatum: physiological and pharmacological characterization. *J. Neurosci.* **12,** 4224–4233 (1992).
15. Gerdeman, G. L., Ronesi, J. & Lovinger, D. M. Postsynaptic endocannabinoid release is critical to long-term depression in the striatum. *Nature Neurosci.* **5,** 446–451 (2002).
16. Kreitzer, A. C. & Malenka, R. C. Dopamine modulation of state-dependent endocannabinoid release and long-term depression in the striatum. *J. Neurosci.* **25,** 10537–10545 (2005).
17. Tang, K., Low, M. J., Grandy, D. K. & Lovinger, D. M. Dopamine-dependent synaptic plasticity in striatum during *in vivo* development. *Proc. Natl Acad. Sci. USA* **98,** 1255–1260 (2001).
18. Calabresi, P. *et al.* Abnormal synaptic plasticity in the striatum of mice lacking dopamine D2 receptors. *J. Neurosci.* **17,** 4536–4544 (1997).
19. Giuffrida, A. *et al.* Dopamine activation of endogenous cannabinoid signaling in dorsal striatum. *Nature Neurosci.* **2,** 358–363 (1999).
20. Betarbet, R., Sherer, T. B. & Greenamyre, J. T. Animal models of Parkinson's disease. *BioEssays* **24,** 308–318 (2002).
21. Schwarting, R. K. & Huston, J. P. Unilateral 6-hydroxydopamine lesions of meso-striatal dopamine neurons and their physiological sequelae. *Prog. Neurobiol.* **49,** 215–266 (1996).
22. Kathuria, S. *et al.* Modulation of anxiety through blockade of anandamide hydrolysis. *Nature Med.* **9,** 76–81 (2003).
23. Di Marzo, V., Hill, M. P., Bisogno, T., Crossman, A. R. & Brotchie, J. M. Enhanced levels of endogenous cannabinoids in the globus pallidus are associated with a reduction in movement in an animal model of Parkinson's disease. *FASEB J.* **14,** 1432–1438 (2000).
24. Gubellini, P. *et al.* Experimental parkinsonism alters endocannabinoid degradation: implications for striatal glutamatergic transmission. *J. Neurosci.* **22,** 6900–6907 (2002).
25. Jenner, P. Dopamine agonists, receptor selectivity and dyskinesia induction in Parkinson's disease. *Curr. Opin. Neurol.* **16** (Suppl. 1), S3–S7 (2003).
26. Makara, J. K. *et al.* Selective inhibition of 2-AG hydrolysis enhances endocannabinoid signaling in hippocampus. *Nature Neurosci.* **8,** 1139–1141 (2005).
27. Saario, S. M. *et al.* URB754 has no effect on the hydrolysis or signaling capacity of 2-AG in the rat brain. *Chem. Biol.* **13,** 811–814 (2006).
28. Baik, J. H. *et al.* Parkinsonian-like locomotor impairment in mice lacking dopamine D2 receptors. *Nature* **377,** 424–428 (1995).
29. Zimmer, A., Zimmer, A. M., Hohmann, A. G., Herkenham, M. & Bonner, T. I. Increased mortality, hypoactivity, and hypoalgesia in cannabinoid CB1 receptor knockout mice. *Proc. Natl Acad. Sci. USA* **96,** 5780–5785 (1999).
30. Compton, D. R., Aceto, M. D., Lowe, J. & Martin, B. R. *In vivo* characterization of a specific cannabinoid receptor antagonist (SR141716A): inhibition of delta 9-tetrahydrocannabinol-induced responses and apparent agonist activity. *J. Pharmacol. Exp. Ther.* **277,** 586–594 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank W. Regehr, B. Sabatini, D. Lovinger, D. Piomelli, J. Surmeier and members of the Malenka laboratory for helpful discussions, X. W. Yang for providing BAC-transgenic mice generated by the GENSAT project, M. Xu-Friedman for providing custom Igor acquisition software, and S. Y. Lee and X. Cai for technical help. This work was supported by a Ruth L. Kirchenstein Fellowship (A.C.K.), a grant from the National Institutes of Health (R.C.M.), and a National Parkinson Foundation Individual Research Grant (R.C.M. and A.C.K.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to R.C.M. (malenka@stanford.edu).

LETTERS

Interleukin-22, a T_H17 cytokine, mediates IL-23-induced dermal inflammation and acanthosis

Yan Zheng[1], Dimitry M. Danilenko[2], Patricia Valdez[1], Ian Kasman[2], Jeffrey Eastham-Anderson[2], Jianfeng Wu[1] & Wenjun Ouyang[1]

Psoriasis is a chronic inflammatory skin disease characterized by hyperplasia of the epidermis (acanthosis), infiltration of leukocytes into both the dermis and epidermis, and dilation and growth of blood vessels[1]. The underlying cause of the epidermal acanthosis in psoriasis is still largely unknown. Recently, interleukin (IL)-23, a cytokine involved in the development of IL-17-producing T helper cells (T_H17 cells)[2,3], was found to have a potential function in the pathogenesis of psoriasis[4,5]. Here we show that IL-22 is preferentially produced by T_H17 cells and mediates the acanthosis induced by IL-23. We found that IL-23 or IL-6 can directly induce the production of IL-22 from both murine and human naive T cells. However, the production of IL-22 and IL-17 from T_H17 cells is differentially regulated. Transforming growth factor-β, although crucial for IL-17 production, actually inhibits IL-22 production. Furthermore, IL-22 mediates IL-23-induced acanthosis and dermal inflammation through the activation of Stat3 (signal transduction and activators of transcription 3) *in vivo*. Our results suggest that T_H17 cells, through the production of both IL-22 and IL-17, might have essential functions in host defence and in the pathogenesis of autoimmune diseases such as psoriasis. IL-22, as an effector cytokine produced by T cells, mediates the cross-talk between the immune system and epithelial cells.

IL-22 belongs to the IL-10 family of cytokines[6]. The IL-22 receptor is expressed on a variety of epithelial tissues[7] and is believed to mediate epithelial innate immunity[8]. IL-22 is produced by activated $CD4^+$ T cells[7] and regulates proliferation and differentiation of keratinocytes[9]. Elevated levels of IL-22 are found in the blood of psoriatic patients[10]. To identify the specific T-cell subset that secretes IL-22, splenocytes from DO11.10 T-cell antigen receptor (TCR) transgenic mice were activated with ovalbumin peptide (OVAp) under various T-cell polarization conditions. We found that IL-22, but not the related cytokines, IL-19, IL-20 or IL-24, was preferentially produced by IL-23-driven T_H17 cells (Supplementary Fig. S1), which is consistent with the results of a recent report[11]. Furthermore, IL-22 could also be induced by IL-23 from activated γδ T cells, $CD8^+$ T cells, and to a smaller extent, monocytes (Supplementary Fig. S1b). T_H17 cells have crucial functions in host defence against infections and have been implicated in the development of autoimmune diseases[12]. To exclude the possibility that IL-23 acted only on memory T cells to produce IL-22 (ref. 13), naive $CD4^+$ T cells were isolated from DO11.10 mice lacking recombination-activating gene 2 ($Rag2^{-/-}$.DO11.10 mice) and were activated by OVAp in the presence of T-cell-depleted splenocytes from Balb/c mice. Again, IL-23-driven T_H17 cells preferentially produced both IL-17 and IL-22 after restimulation (Fig. 1a). Moreover, the addition of either interferon (IFN)-γ or IL-4 completely abolished IL-17 production[3], but could only moderately inhibit IL-22 production (Fig. 1b). In contrast, when fully established T_H17 cells were re-activated, neither IFN-γ nor IL-4 was able to reprogramme the phenotype of these cells and inhibit the production of either IL-17 or IL-22 (Fig. 1c). Furthermore, when these cells were restimulated weekly with OVAp and antigen-presenting

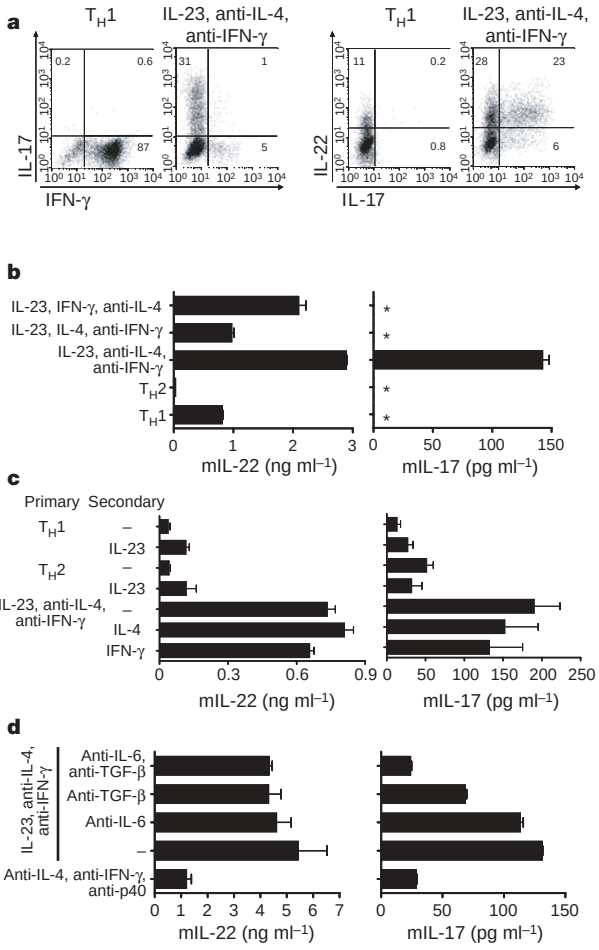


Figure 1 | IL-22 is a new effector cytokine downstream of IL-23. $CD4^+$ T cells from $Rag2^{-/-}$.DO11.10 mice were stimulated under the indicated conditions for 4 days. **a**, Intracellular cytokine staining for IFN-γ, IL-17 and IL-22 on day 3. Plots are gated on $CD4^+$ T cells. **b–d**, Supernatant was collected 72 h after activation. Mouse (m)IL-22 and mIL-17 production in the culture supernatant were measured by ELISA; results are reported as means ± s.d. Asterisk, below the detection limit for ELISA. **c**, For secondary activation, primary activated T cells from **b** were rested for two days after expansion and restimulated with anti-CD3 and anti-CD28 mAbs. All data are representative of two independent experiments.

[1]Department of Immunology and [2]Department of Pathology, Genentech, Inc., South San Francisco, California 94080, USA.

cells (APCs) for a month without further addition of exogenous IL-23, they maintained their ability to produce both IL-17 and IL-22 (data not shown). These data suggest that this lineage is stable.

IL-6 and transforming growth factor (TGF)-β were both required for the induction of T_H17 cells, whereas IL-23 was necessary for the establishment of the T_H17 lineage[14–17]. A recent publication suggested a similar regulation for murine IL-22 production[11]. To clarify whether IL-6 and TGF-β were necessary for the IL-23-stimulated production of IL-22 in the OVAp/APC system, endogenous IL-6 and TGF-β were neutralized by blocking antibodies. As expected, blocking IL-6 and TGF-β completely inhibited the IL-23-induced production of IL-17 (Fig. 1d). However, IL-22 production was mostly sustained, indicating that IL-6 and TGF-β were dispensable for the IL-23-induced production of IL-22. To exclude the possibility that other factors from APCs might contribute to the regulation of IL-22 and IL-17, naive $CD4^+$ T cells (more than 98%) were activated by anti-CD3 and anti-CD28 monoclonal antibodies (mAbs) under various conditions (Fig. 2). We verified that IL-23, TGF-β or IL-6 alone did not induce significant IL-17 production[15], whereas IL-6 plus TGF-β promoted T_H17 differentiation (Fig. 2a). TGF-β alone failed to promote IL-22 production. IL-23 on its own again induced IL-22 production under APC-free conditions (Fig. 2a), and this induction was independent of IL-6 because similar results were also obtained with IL-$6^{-/-}$ T cells (Fig. 2c). Unexpectedly, IL-6 alone also induced IL-22 production in both this APC-free system (Fig. 2a) and the OVAp/APC system (Supplementary Fig. S1d). TGF-β inhibited the IL-6-induced production of IL-22 in a dose-dependent manner (Fig. 2b, left panel). At doses of TGF-β greater than 10 ng ml^{-1}, only trace amounts of IL-22 were detected (Fig. 2a, b). In contrast, TGF-β promoted IL-17 production at lower concentrations; maximal IL-17 production was attained at 1 ng ml^{-1} TGF-β (Fig. 2b, right panel). Even at a very high concentration of TGF-β (200 ng ml^{-1}), T cells still produced significant amounts of IL-17 (more than 10 ng ml^{-1}). These data indicate that although IL-22 and IL-17 can be produced by the same T_H17 cells, regulation of their production at the molecular level may be distinct. Furthermore, studying the regulation of human T_H17 cells, we observed that either IL-23 or IL-6 primed naive human $CD4^+$ T cells to differentiate into IL-22-producing cells (Supplementary Fig. S2). However, none of the conditions reported to prime murine T_H17 polarization induced significant human IL-17 production (less than 9 pg ml^{-1}; data not shown). Thus, our data clearly show that either IL-23 or IL-6 can induce *de novo* IL-22 production from both murine and human naive T cells.

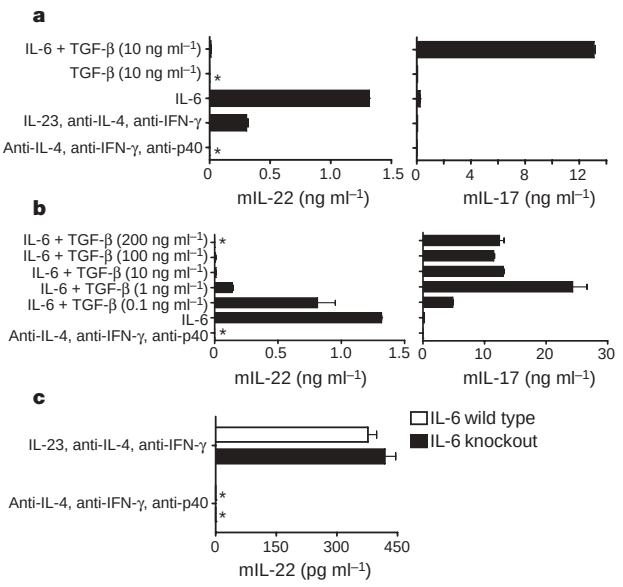


Figure 2 | Differential regulation of IL-22 and IL-17 production from $CD4^+$ T cells. FACS-sorted naive $CD4^+$ T cells from C57Bl/6 mice (**a**–**c**) and IL-$6^{-/-}$ mice (**c**) were stimulated with anti-CD3 and anti-CD28 mAbs in the presence of the indicated cytokine and/or neutralizing antibody for 72 h. For all experiments, culture supernatant was collected after 72 h, then mIL-22 and mIL-17 production in the culture supernatant were measured by ELISA; results are reported as means ± s.d. Asterisk, below detection limit for ELISA. Data are representative of two (**b**, **c**) or three (**a**) independent experiments.

Overexpression of IL-23, but not IL-12, has been detected in lesional psoriatic skin and in a murine model of inflammatory skin disease[4,18]. When injected into mouse ear, however, either IL-12 or IL-23 induced an increase in ear thickness (Supplementary Fig. S3a), similarly to a previous report[19]. Histological analysis revealed that ears injected with either IL-12 or IL-23 developed epidermal acanthosis with inflammatory cellular infiltration (Fig. 3a, and Supplementary Fig. S3b). However, whereas IL-12 induced mild to moderate acanthosis with a marked, predominantly mononuclear, dermal inflammatory cellular infiltration (Fig. 3a), IL-23 induced marked acanthosis with a mixed dermal inflammatory cellular infiltration that also included increased numbers of neutrophils and eosinophils (Fig. 3a). FACS (fluorescence-activated cell sorting) analysis of the infiltrating lymphocytes demonstrated the presence of both $CD4^+$ and $CD8^+$ T cells in the inflamed dermis of both the IL-12-treated and IL-23-treated mice (Supplementary Fig. S3c). Immunohistochemical staining revealed markedly increased Stat3 activation in epidermal keratinocytes with IL-23 treatment in comparison with treatment with either PBS or IL-12 (Fig. 3b). Reverse transcriptase polymerase chain reaction (RT–PCR) analysis showed that IL-12 induced a significant increase in IFN-γ expression, whereas IL-23 induced the production of both IL-22 and IL-17 after injection into the ear (Fig. 3c). To further confirm that these cytokines

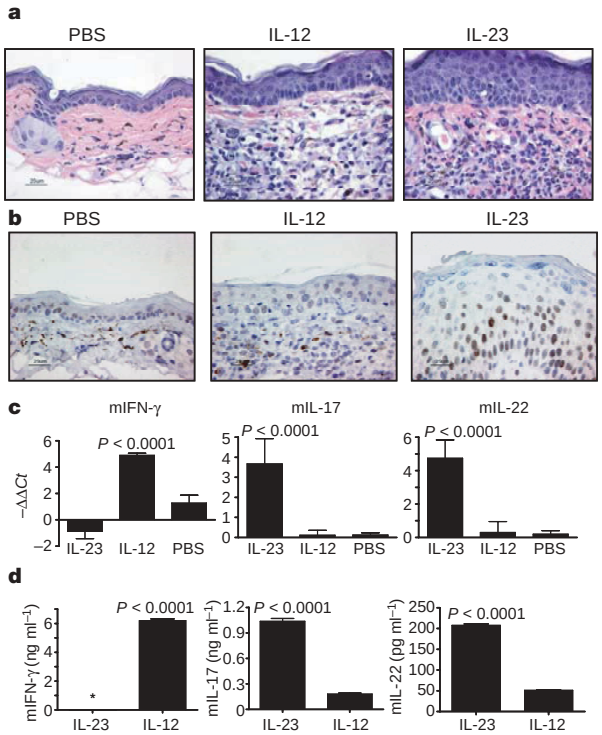


Figure 3 | IL-12 and IL-23 both induce acanthosis and inflammation in mouse ear skin, but elicit distinct cytokine profiles. Ears from C57Bl/6 mice ($n = 5$ for each group) were each injected intradermally every other day with 500 ng of IL-12, 500 ng of IL-23 or PBS in a total volume of 20 μl. **a**, **b**, On day 16, ears were collected for staining with H&E (**a**) and pStat3 immunohistochemical staining (**b**). Scale bar, 25 μm. **c**, On day 8, real-time RT–PCR was performed on RNA isolated from individual treated mouse ears from each group. **d**, On day 8, infiltrating cells in treated ears were also restimulated with anti-CD3 and anti-CD28 mAbs for 72 h. Supernatant was collected for cytokine ELISA and results are reported as means ± s.d. Asterisk, below detection limit for ELISA. All data are representative of two independent experiments.

were produced by the lymphocytes that infiltrated into the ears, we collected lymphocytes migrating from treated ears and measured the cytokine production after activation. Consistently, T cells from IL-23-treated ears preferentially produced IL-22 and IL-17, whereas T cells from IL-12-treated ears secreted IFN-γ after stimulation with anti-CD3 and anti-CD28 (Fig. 3d). $CD4^+$ T cells from IL-23-treated ears produced IL-17 and IL-22 simultaneously, as shown by intracellular staining for cytokines. In contrast, $CD4^+$ T cells from IL-12-treated ears expressed predominantly IFN-γ (Supplementary Fig. S3d). Together, these *in vivo* data corroborated our *in vitro* observations that IL-23 promotes the production of IL-22 and IL-17 from T cells.

IL-22-deficient mice were generated to further address the role of IL-22 in IL-23-induced acanthosis and dermal inflammation (Supplementary Fig. S4a, b). These mice showed normal development with similar lymphocyte composition and development in all major lymphoid organs when compared with wild-type mice (data not shown). Activation and differentiation of IL-$22^{-/-}$ $CD4^+$ T cells was similar to that of their wild-type counterparts, with the exception of IL-22 production (Supplementary Fig. S4c, d, and data not shown). IL-23-induced ear swelling was significantly decreased in IL-$22^{-/-}$ mice in comparison with the control groups (Fig. 4a). Both epidermal acanthosis (Fig. 4b) and dermal inflammation were significantly decreased in the ears of IL-$22^{-/-}$ mice (Fig. 4d) compared with IL-23-treated wild-type littermates (Fig. 4c). However, IL-$22^{-/-}$ mouse ears still showed a minimal degree of epidermal hyperplasia and dermal inflammation compared with PBS-treated wild-type mouse ears (Fig. 4e). We did not detect elevated IL-19, IL-20 or IL-24, all of which use IL-20 receptor-β (IL-20Rβ) to signal, in response to either IL-12 or IL-23 injection (Supplementary Figs S1b and S5). However, we cannot exclude the possibility that signalling through IL-20Rβ might mediate the residual epidermal hyperplasia, as has recently been suggested[20]. IL-17 may also have a potential synergistic effect on keratinocyte hyperplasia in this model, because IL-17 has been shown to induce the expression of proinflammatory proteins from human keratinocytes, particularly in combination with other cytokines such as IFN-γ (ref. 21). Furthermore, the different phenotypes we observed in IL-$22^{-/-}$ mice after injection of IL-23 were not due to any developmental defects in this strain, because similar decreases in ear thickness, acanthosis and dermal inflammation were also observed in wild-type mice receiving a neutralizing anti-IL-22 mAb in comparison with those receiving an isotype control mAb (Supplementary Fig. S6). In contrast to its effect on IL-23-injected ears, IL-22 blockade or IL-22 deficiency had no effect at all on IL-12-induced epidermal acanthosis or dermal inflammation (Supplementary Fig. S7). These results are consistent with the observation that IL-23, but not IL-12, promotes IL-22 production *in vitro* and *in vivo*, and that activation of Stat3 from epidermal keratinocytes is markedly induced by IL-23 but not by IL-12 (Fig. 3b).

To understand better the downstream mechanisms that mediate IL-22-induced epidermal hyperplasia, we analysed the expression of keratin 6 and activated phospho-Stat3 (pStat3), both of which are elevated in psoriatic skin and are markers for keratinocyte proliferation[22,23]. Although fewer keratinocytes stained positive for keratin 6 in the IL-$22^{-/-}$ mice than in the wild-type mice, this difference was attributable to a decrease in total epidermal thickness and not to any difference in staining intensity (Supplementary Fig. S8c). Many other factors, such as IL-1 and epidermal growth factor (EGF)[24,25], can upregulate the epidermal expression of keratin 6, and it is possible that one or more of these factors contributed to the increased epidermal expression of keratin 6 in the IL-$22^{-/-}$ mice after administration of IL-23. In contrast, we observed significantly less activation of pStat3 in IL-$22^{-/-}$ mice treated with IL-23 than in mice in the wild-type control group (Fig. 4f–h). These data indicate that much of the IL-23-induced epidermal acanthosis may be mediated through Stat3 activation in keratinocytes, a pathway that is critical for the development of psoriasis[22]. We also scored the dermal inflammatory infiltration histologically with immunohistochemical staining to identify different leukocyte populations, and showed that IL-23-treated IL-$22^{-/-}$ mice had reduced neutrophilic, but similar $CD3^+$ T-cell, infiltration compared with that in wild-type mice (Supplementary Fig. S8a, b). The decreased inflammatory cellular infiltration, especially that of neutrophils, evident in IL-22-deficient/antibody-blocked ears, was probably an indirect effect as a result of reduced stimulation of keratinocytes in the absence of IL-22. This hypothesis is further supported by the facts that immune cells do not express

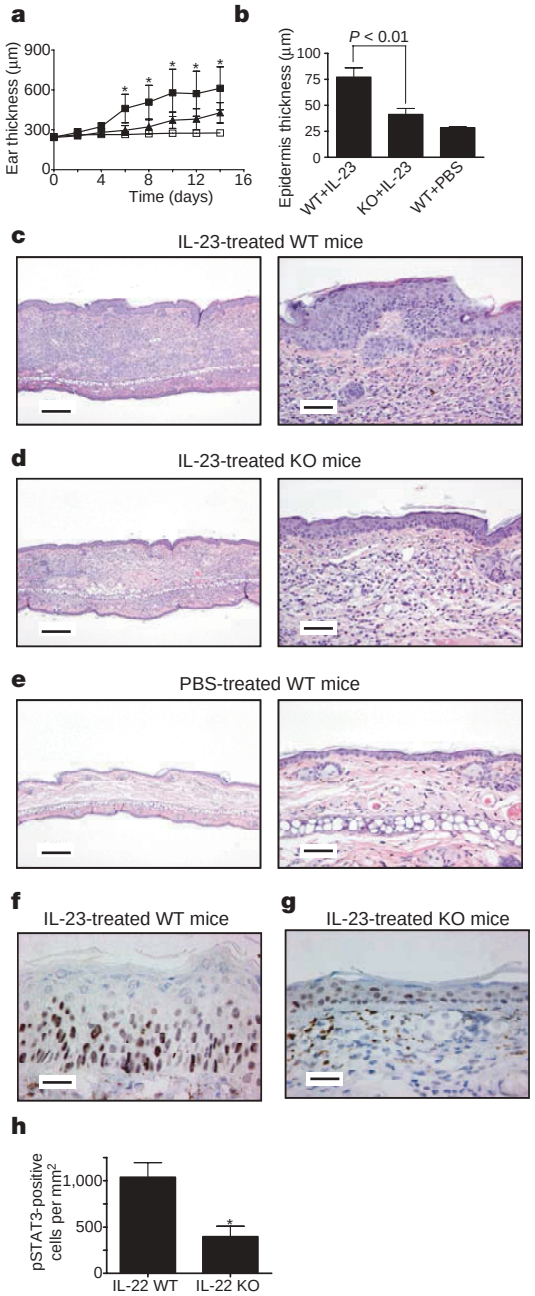


Figure 4 | IL-22 deficiency decreases IL-23-induced ear skin acanthosis and inflammation. Ears from IL-$22^{-/-}$ ($n = 5$) and wild-type (WT) littermate ($n = 5$) mice were each injected intradermally every other day with 500 ng of IL-23 or PBS in a total volume of 20 μl. **a**, Ear thickness was measured before and at multiple time points after injection; results are reported as means ± s.d. Asterisk, $P < 0.001$. Filled squares, WT plus IL-23; triangles, knockout (KO) plus IL-23; open squares, WT plus PBS. **b–h**, On day 14, ears were collected for H&E staining (**b–e**) and pStat3 immunohistochemical staining (**f–h**) and analysed for average epidermal thickness (**b**) and pStat3-positive cells per mm^2 of epidermis (**h**). Scale bars, 200 μm (left panels of **c–e**), 50 μm (right panels of **c–e**); 25 μm (**f, g**). Asterisk, $P < 0.001$. All data are representative of two independent experiments.

IL-22 receptors, and IL-22 can induce chemokine production from various cell lines, including keratinocytes[9,10]. Taken together, our data suggest that IL-22 has a crucial function in the development of dermal inflammation and epidermal acanthosis induced by IL-23, but not that induced by IL-12.

Psoriasis is a complex autoimmune disease, and the interactions between dendritic cells (DCs), T cells and keratinocytes provide the basis for the unique molecular signature defining a typical psoriatic lesion[26–28]. Although the epidermal acanthosis and dermal inflammation induced by IL-23 injection into the ear does not model psoriasis, many of the features in this model, such as IL-22 upregulation, epidermal hyperplasia and pStat3 activation, are similar to features evident in psoriasis. Our studies suggest a potential molecular mechanism for the epidermal hyperplasia resulting from leukocyte infiltration and inflammation in psoriasis. This recently characterized T_H17 pathway, in addition to the T_H1 pathway[29,30], may provide new therapeutic targets for the treatment of psoriasis.

METHODS

Mice. C57Bl/6 and IL-$6^{-/-}$ mice were purchased from the Jackson Laboratory. Balb/c mice were purchased from Charles River Laboratories. DO11.10 TCR transgenic and *Rag2*$^{-/-}$ DO11.10 TCR transgenic mice were bred in our facility. IL-$22^{-/-}$ mice were generated by Lexicon Genetics Incorporated by using strategies as described in Supplementary Methods.

Cell purifications. $CD4^+$ T cells, $CD8^+$ T cells, γδ T cells, monocytes and naive $CD4^+$ T cells were purified from mouse spleens as described in Supplementary Methods.

***In vitro* cell culture conditions.** $CD4^+$ T cells from *Rag2*$^{-/-}$ DO11.10 TCR transgenic mice were activated with OVAp in the presence of irradiated splenocytes. For activation of naive T cells, cells were stimulated with anti-CD3 and anti-CD28 mAbs. Where indicated, cultures were supplemented with recombinant cytokines and/or antibodies as described in Supplementary Methods.

RNA isolation and real-time RT–PCR. Cell and tissue RNA were isolated with an RNeasy Mini Kit (Qiagen). Real-time RT–PCR was conducted on an ABI 7500 Real-Time PCR system (Applied Biosystems) with primers and probes as described in Supplementary Methods.

Cytokine production and intracellular cytokine staining. Cytokine production was determined by enzyme-linked immunosorbent assay (ELISA) as described in Supplementary Methods. For intracellular staining, cells were stimulated for 6 h with 50 ng ml^{-1} 12-*O*-tetradecanoylphorbol-13-acetate (Sigma) and 750 ng ml^{-1} ionomycin (Calbiochem) or not at all. After 1 h, Brefeldin A (Epicentre Biotechnologies) was added to block cytokine secretion for a further 5 h. Cells were stained as described in Supplementary Methods.

MetaMorph software analysis of ear skin epidermis thickness. Haematoxylin/eosin (H&E)-stained ear skin images were acquired and analysed by MetaMorph (Molecular Devices) as described in Supplementary Methods.

Immunohistochemistry for pStat3 and image analysis. Sections of mouse ear were stained for pStat3, and images from pStat3-stained slides were acquired by the Ariol SL-50 automated image analysis system (Applied Imaging), and were then analysed using Ariol software as described in Supplementary Methods.

Received 5 November; accepted 6 December 2006.
Published online 24 December 2006.

1. Nickoloff, B. J. & Nestle, F. O. Recent insights into the immunopathogenesis of psoriasis provide new therapeutic opportunities. *J. Clin. Invest.* **113,** 1664–1675 (2004).
2. Park, H. *et al.* A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. *Nature Immunol.* **6,** 1133–1141 (2005).
3. Harrington, L. E. *et al.* Interleukin 17-producing $CD4^+$ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nature Immunol.* **6,** 1123–1132 (2005).
4. Lee, E. *et al.* Increased expression of interleukin 23 p19 and p40 in lesional skin of patients with psoriasis vulgaris. *J. Exp. Med.* **199,** 125–130 (2004).
5. Kauffman, C. L. *et al.* A phase I study evaluating the safety, pharmacokinetics, and clinical response of a human IL-12 p40 antibody in subjects with plaque psoriasis. *J. Invest. Dermatol.* **123,** 1037–1044 (2004).
6. Pestka, S. *et al.* Interleukin-10 and related cytokines and receptors. *Annu. Rev. Immunol.* **22,** 929–979 (2004).
7. Gurney, A. L. IL-22, a Th1 cytokine that targets the pancreas and select other peripheral tissues. *Int. Immunopharmacol.* **4,** 669–677 (2004).
8. Wolk, K. *et al.* IL-22 increases the innate immunity of tissues. *Immunity* **21,** 241–254 (2004).
9. Boniface, K. *et al.* IL-22 inhibits epidermal differentiation and induces proinflammatory gene expression and migration of human keratinocytes. *J. Immunol.* **174,** 3695–3702 (2005).
10. Wolk, K. *et al.* IL-22 regulates the expression of genes responsible for antimicrobial defense, cellular differentiation, and mobility in keratinocytes: a potential role in psoriasis. *Eur. J. Immunol.* **36,** 1309–1323 (2006).
11. Liang, S. C. *et al.* Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *J. Exp. Med.* **203,** 2271–2279 (2006).
12. McKenzie, B. S., Kastelein, R. A. & Cua, D. J. Understanding the IL-23-IL-17 immune pathway. *Trends Immunol.* **27,** 17–23 (2006).
13. Parham, C. *et al.* A receptor for the heterodimeric cytokine IL-23 is composed of IL-12Rβ1 and a novel cytokine receptor subunit, IL-23R. *J. Immunol.* **168,** 5699–5708 (2002).
14. Weaver, C. T., Harrington, L. E., Mangan, P. R., Gavrieli, M. & Murphy, K. M. Th17: an effector CD4 T cell lineage with regulatory T cell ties. *Immunity* **24,** 677–688 (2006).
15. Veldhoen, M., Hocking, R. J., Atkins, C. J., Locksley, R. M. & Stockinger, B. TGFβ in the context of an inflammatory cytokine milieu supports de novo differentiation of IL-17-producing T cells. *Immunity* **24,** 179–189 (2006).
16. Mangan, P. R. *et al.* Transforming growth factor-β induces development of the T_H17 lineage. *Nature* **441,** 231–234 (2006).
17. Bettelli, E. *et al.* Reciprocal developmental pathways for the generation of pathogenic effector T_H17 and regulatory T cells. *Nature* **441,** 235–238 (2006).
18. Kopp, T. *et al.* IL-23 production by cosecretion of endogenous p19 and transgenic p40 in keratin 14/p40 transgenic mice: evidence for enhanced cutaneous immunity. *J. Immunol.* **170,** 5438–5444 (2003).
19. Kopp, T. *et al.* Inflammatory skin disease in K14/p40 transgenic mice: evidence for interleukin-12-like activities of p40. *J. Invest. Dermatol.* **117,** 618–626 (2001).
20. Chan, J. R. *et al.* IL-23 stimulates epidermal hyperplasia via TNF and IL-20R2-dependent mechanisms with implications for psoriasis pathogenesis. *J. Exp. Med.* **203,** 2577–2587 (2006).
21. Teunissen, M. B., Koomen, C. W., de Waal Malefyt, R., Wierenga, E. A. & Bos, J. D. Interleukin-17 and interferon-γ synergize in the enhancement of proinflammatory cytokine production by human keratinocytes. *J. Invest. Dermatol.* **111,** 645–649 (1998).
22. Sano, S. *et al.* Stat3 links activated keratinocytes and immunocytes required for development of psoriasis in a novel transgenic mouse model. *Nature Med.* **11,** 43–49 (2005).
23. Stoler, A., Kopan, R., Duvic, M. & Fuchs, E. Use of monospecific antisera and cRNA probes to localize the major changes in keratin expression during normal and abnormal epidermal differentiation. *J. Cell Biol.* **107,** 427–446 (1988).
24. Komine, M. *et al.* Interleukin-1 induces transcription of keratin K6 in human epidermal keratinocytes. *J. Invest. Dermatol.* **116,** 330–338 (2001).
25. Jiang, C. *et al.* Epidermal growth factor and transforming growth factorα specifically induce the activation- and hyperproliferation-associated keratins 6 and 16. *Proc. Natl Acad. Sci. USA* **90,** 6786–6790 (1993).
26. Lowes, M. A. *et al.* Increase in TNF-α and inducible nitric oxide synthase-expressing dendritic cells in psoriasis and reduction with efalizumab (anti-CD11a). *Proc. Natl Acad. Sci. USA* **102,** 19057–19062 (2005).
27. Abrams, J. R. *et al.* Blockade of T lymphocyte costimulation with cytotoxic T lymphocyte-associated antigen 4-immunoglobulin (CTLA4Ig) reverses the cellular pathology of psoriatic plaques, including the activation of keratinocytes, dendritic cells, and endothelial cells. *J. Exp. Med.* **192,** 681–694 (2000).
28. Nestle, F. O., Turka, L. A. & Nickoloff, B. J. Characterization of dermal dendritic cells in psoriasis. Autostimulation of T lymphocytes and induction of Th1 type cytokines. *J. Clin. Invest.* **94,** 202–209 (1994).
29. Austin, L. M., Ozawa, M., Kikuchi, T., Walters, I. B. & Krueger, J. G. The majority of epidermal T cells in psoriasis vulgaris lesions can produce type 1 cytokines, interferon-γ, interleukin-2, and tumor necrosis factor-α, defining TC1 (cytotoxic T lymphocyte) and TH1 effector populations: a type 1 differentiation bias is also measured in circulating blood T cells in psoriatic patients. *J. Invest. Dermatol.* **113,** 752–759 (1999).
30. Schlaak, J. F. *et al.* T cells involved in psoriasis vulgaris belong to the Th1 subset. *J. Invest. Dermatol.* **102,** 145–149 (1994).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Chan, H. Spits, M. Townsend, J. Grogan and B. Irving for critical suggestions, and W. Lee, C. Olsson, J. Starks, L. Gilmour, M. Hamilton, L. Hall and the Genentech Histology and Immunohistochemistry Laboratories for technical assistance.

Author Contributions Y.Z. performed most of the experimental work. D.M.D., I.K. and J.E.-A. provided all the histology data analyses. P.V. conducted the *in vitro* experiments with mouse monocytes and human T cells. J.W. developed the assays for IL-22 ELISA and intracellular staining, and genotyped the IL-22 knockout mice. W.O. devised and planned the project, and the manuscript was written by Y.Z., D.M.D. and W.O.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the paper on www.nature.com/nature. Correspondence and requests for materials should be addressed to W.O. (ouyang.wenjun@gene.com).

LETTERS

Direct control of shoot meristem activity by a cytokinin-activating enzyme

Takashi Kurakawa[1]*, Nanae Ueda[2]*, Masahiko Maekawa[3], Kaoru Kobayashi[1], Mikiko Kojima[2], Yasuo Nagato[1], Hitoshi Sakakibara[2] & Junko Kyozuka[1]

The growth of plants depends on continuous function of the meristems. Shoot meristems are responsible for all the post-embryonic aerial organs, such as leaves, stems and flowers[1]. It has been assumed that the phytohormone cytokinin has a positive role in shoot meristem function[2–4]. A severe reduction in the size of meristems in a mutant that is defective in all of its cytokinin receptors has provided compelling evidence that cytokinin is required for meristem activity[5,6]. Here, we report a novel regulation of meristem activity, which is executed by the meristem-specific activation of cytokinins. The *LONELY GUY (LOG)* gene of rice is required to maintain meristem activity and its loss of function causes premature termination of the shoot meristem. *LOG* encodes a novel cytokinin-activating enzyme that works in the final step of bioactive cytokinin synthesis. Revising the long-held idea of multistep reactions, LOG directly converts inactive cytokinin nucleotides to the free-base forms, which are biologically active, by its cytokinin-specific phosphoribohydrolase activity. *LOG* messenger RNA is specifically localized in shoot meristem tips, indicating the activation of cytokinins in a specific developmental domain. We propose the fine-tuning of concentrations and the spatial distribution of bioactive cytokinins by a cytokinin-activating enzyme as a mechanism that regulates meristem activity.

Cytokinins are N^6-substituted adenine derivatives that have a crucial role in many aspects of plant growth and development[7,8]. Like the other plant hormones, cytokinin activity *in planta* is thought to be controlled by a balance of synthesis, catabolism and inactivating conjugations[9,10]. Moreover, the regulation of cytokinin signalling further modulates its effects[11,12]. The spatial and temporal distribution of bioactive cytokinin levels is also strictly controlled in plant development events. Recent studies have shown that cytokinins are synthesized and catalysed at various sites in the plant[13–15], but the exact mechanism whereby cytokinin activity is regulated is not entirely clear.

The rice mutant *lonely guy (log)* was identified in a screen for defects in the maintenance of shoot meristems. In addition to a severe reduction of the panicle size and abnormal branching patterns (Fig. 1a), the number of floral organs was also decreased in *log* mutants (Fig. 1b–e), and flowers often contained only one stamen but no pistil (thus '*lonely guy*'). The inner floral organs were affected more severely than the outer organs in all allelic mutants examined (Fig. 1e). We view these phenotypes as the result of defects in maintaining meristem activity and have tested this idea by examining the vegetative and reproductive meristems in detail. Although no clear abnormality was observed during vegetative development, the shoot apical meristem (SAM) of *log-1* was smaller than that of wild type (Supplementary Table 1). After transition to the reproductive phase, a number of meristems that give rise to panicle branches and flowers are generated in the wild-type inflorescence. However, *log* mutant inflorescence meristems and panicle branch meristems abort soon after the production of a few lateral meristems, leading to the production of a small panicle with a reduced number of branches and flowers (Fig. 1a). The wild-type floral meristem continues organ differentiation until it generates a pistil (Fig. 1f), but the *log* mutant

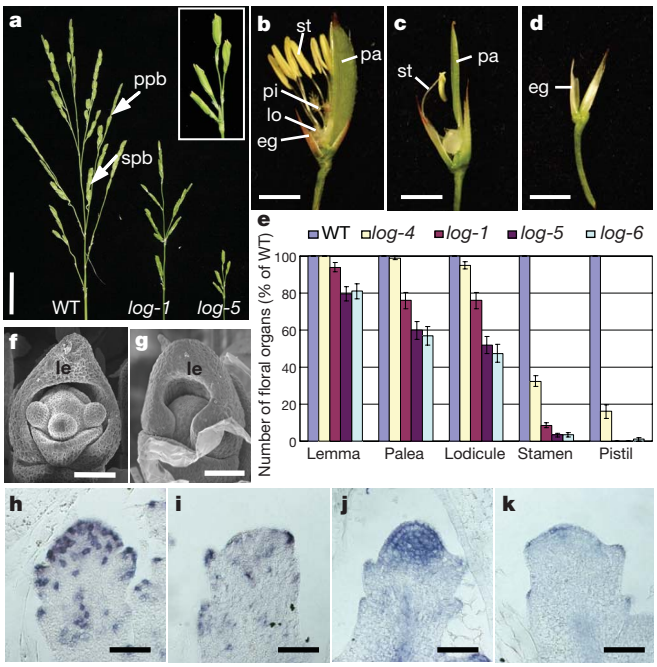


Figure 1 | *log* mutant phenotypes. **a**, Panicles of wild type (left), *log-1* (centre) and *log-5* (right). The *log* panicles are small and show altered branching patterns. The inset shows a closer view of the *log-5* panicle. **b**, Wild-type flower. **c**, *log-1* flower with a weak phenotype, which has only one stamen and no pistil. In **b** and **c**, the lemma is removed to show the inner floral organs. **d**, *log-1* flower with a strong phenotype, which has only a pair of empty glumes. **e**, Reduction of the floral organ number in *log* mutants. The reduction in floral organ number in *log* mutants is given as per cent (mean ± s.d.) of wild type ($n = 100$). **f**, **g**, Scanning electron micrograph of an immature flower of wild type (**f**) and *log-1* (**g**) at the carpel initiation stage. Three stamen primordia are seen in the wild-type flower (**f**). **h**–**k**, Comparison of marker gene expression in wild type (**h**, **j**) and *log-1* (**i**, **k**); *in situ* hybridization of histone *H4* (**h**, **i**) or *OSH1* (**j**, **k**) expression. ppb, primary panicle branch; spb, secondary panicle branch; eg, empty glume; le, lemma; lo, lodicule; pa, palea; pi, pistil, st, stamen. Scale bars, 3 cm (**a**); 3 mm (**b**, **c**, **d**); and 50 μm (**f** to **k**).

[1]Graduate School of Agricultural and Life Sciences, University of Tokyo, Yayoi, Bunkyo, Tokyo 113-8652, Japan. [2]RIKEN Plant Science Center, Tsurumi, Yokohama 230-0045, Japan. [3]Research Institute for Bioresources, Okayama University, Kurashiki, Okayama, 710-0046 Japan.
*These two authors contributed equally to this work.

floral meristem flattens and stops organ differentiation prematurely, without forming the normal number of floral organs (Fig. 1g). In a wild-type floral meristem, many cells express histone *H4*, accompanying the differentiation of floral organs (Fig. 1h). In contrast, very few *log* mutant cells express histone *H4* (Fig. 1i). *OSH1*, a meristem marker encoding a KNOTTED1-type homeobox (KNOX) protein, is strongly expressed in all wild-type floral meristems[16] (Fig. 1j), whereas it is expressed only in the upper region of the flat meristem of *log*, indicating that few meristematic cells remain (Fig. 1k). These results strongly support the notion that meristem activity is not properly maintained in *log* mutants.

The *LOG* gene was isolated by positional cloning. Fine mapping using 442 mutant F_2 plants delimited the gene to a 35 kb region on chromosome 1, in which there were four putative genes (Fig. 2a). Mutations in one of the four predicted genes were found in all six *log* mutant alleles, whereas no mutation was found in any of the other three predicted genes (Fig. 2b; Supplementary Fig. 1). The mutant phenotype of *log-2* was complemented by the introduction of a full-length complementary DNA driven by a 1.5 kb promoter region of *LOG* (Fig. 2c, d). Thus, on the basis of these results, we concluded that the *LOG* gene had been isolated. This gene (LOC_Os01g40630) consists of 7 exons and 6 introns and encodes a polypeptide composed of 242 amino acids (Fig. 2b). There are 10 genes that have high sequence similarity to *LOG* in the rice genome, and *Arabidopsis* has a similar number of homologues (Supplementary Fig. 2). The 11 LOG family genes show a variety of expression patterns (Supplementary Fig. 3).

LOG mRNA expression was examined by *in situ* hybridization analysis. *LOG* is expressed weakly in a small area in the upper part of the SAM and axillary meristems of vegetative shoot apices (Fig. 3a–c). Interestingly, the absence of *LOG* transcripts from the site where primordium differentiation for the next leaf will take place was often observed. A higher level of *LOG* expression was repeatedly observed in all meristems of the developing panicle (Fig. 3d), including the panicle branch (Fig. 3e) and floral meristems (Fig. 3f). The strong *LOG* signal is localized in two or three layers of cells at the tip of the meristem and is not observed in the rib domain. To further define where *LOG* is expressed, the mRNA distributions of *LOG*, *OSH1* and *OsCLV3*, a putative *CLAVATA 3* orthologue, were compared (Supplementary Fig. 4). This comparison clearly indicates that *LOG* expression is restricted to a subset of cells in the meristem, and includes the entire region of presumptive stem cells. Therefore, we propose that *LOG* expression defines a new region in the meristem and that this region may play a novel part in the maintenance of meristem activity.

LOG is annotated as a lysine decarboxylase, which catalyses the decarboxylation of L-lysine to generate cadaverine, a kind of polyamine (http://rapdb.dna.affrc.go.jp/). However, we were not able to detect lysine decarboxylase activity with purified recombinant LOG protein (data not shown). Genes homologous to *LOG* are found in a wide range of organisms including bacteria, but their functions have not been assigned. Curiously, in some soil phytopathogenic bacteria such as *Agrobacterium rhizogenes*[17] and *Rhodococcus fascians*[18], the *LOG* homologues (*riorf52* and *fas* ORF6, respectively) are located adjacent to the gene for adenosine phosphate-isopentenyltransferase (IPT), which catalyses the initial step of cytokinin synthesis. This led us to pursue the possibility that LOG is involved in the control of cytokinin metabolism. The first steps in plant cytokinin synthesis result in the production of N^6-(Δ^2-isopentenyl)adenine (iP) riboside 5′-tri-, di- or monophosphate (iPRTP, iPRDP or iPRMP, respectively) by IPT; these can then be hydroxylated to *trans*-zeatin (tZ) riboside 5′-tri-, di- or mono-phosphate (tZRTP, tZRDP or tZRMP, respectively) by a cytochrome P450 monooxygenase, CYP735A[10] (Fig. 4a). Active cytokinin species are free-base forms, such as iP and tZ. The current model is that cytokinin nucleotides are converted to active forms in a two-step reaction catalysed by nucleotidase[19] and nucleosidase[20], but the responsible genes have not been identified (Fig. 4a). Unexpectedly, we found that LOG has phosphoribohydrolase activity, which directly converts a cytokinin nucleotide such as iPRMP and tZRMP to the free-base form with the release of a ribose 5′-monophosphate (Fig. 4b). LOG specifically reacts with cytokinin nucleoside 5′-monophosphates, but not with the di- or triphosphate, AMP (Fig. 4c), or any cytokinin ribosides or bases (data not shown). Dephosphoribosylation of the cytokinin nucleotide by LOG thus represents a new cytokinin-activating reaction. The K_m value of LOG for iPRMP is 11.7 μM, and for tZRMP is 22.0 μM; specific activity is 5.6 and 4.2 μmol min^{-1} mg^{-1} protein, respectively. Transient expression of *LOG–GFP* shows that LOG functions in the cytosol (Fig. 4d; Supplementary Fig. 5). Cytokinin nucleotide phosphoribohydrolase activity has been confirmed for some of the proteins encoded by the *Arabidopsis LOG* homologues (T. Kuroha and H.S., unpublished).

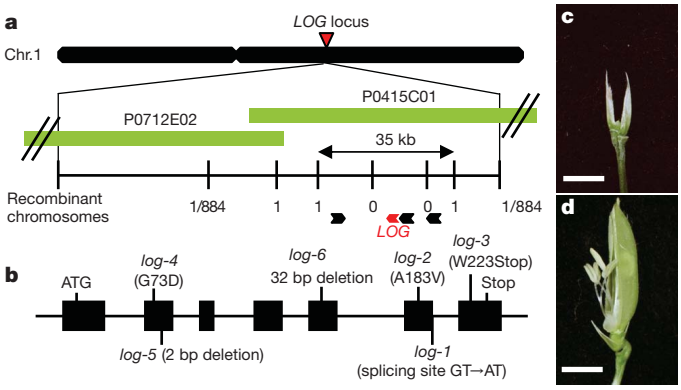


Figure 2 | Positional cloning of *LOG*. **a**, Fine mapping of the *LOG* locus. The region of the *LOG* locus was narrowed within a 35 kb region on chromosome 1 (Chr. 1), which contained 4 predicted genes. **b**, Exon/intron structure of *LOG*. The *LOG* gene contains 7 exons (black box) and 6 introns. Four mutant alleles of the *LOG* gene contain a base substitution that causes an alteration of the splice site (*log-1*), or an amino acid exchange (*log-2* and *-4*), or the generation of a premature stop codon (*log-3*). Two mutant alleles, *log-5* and *log-6*, contain a deletion of 2 and 32 base pairs, respectively. **c**, Complementation test. Defects in *log-2* (**c**) were rescued by the introduction of *LOG* cDNA driven by the *LOG* gene promoter (**d**). Scale bars, 3 mm.

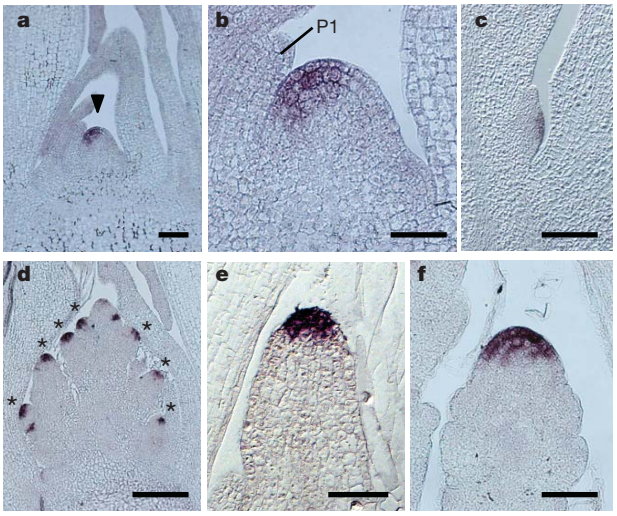


Figure 3 | Expression pattern of *LOG* mRNA. **a**, **b**, *LOG* expression in the SAM of wild-type meristems. P1 indicates a young leaf at the P1 stage. **c**, *LOG* expression in the axillary meristem. **d**, *LOG* expression in a developing panicle at the primary branch initiation stage. **e**, **f**, A closer view of *LOG* expression in a primary panicle branch meristem (**e**) and a floral meristem (**f**). The arrowhead in (**a**) indicates the SAM. Asterisks in (**d**) indicate primary or secondary panicle branch meristems. Scale bars, 100 μm (**a**); 50 μm (**b**, **c**, **e**, **f**); and 200 μm (**d**).

The newly identified molecular function of LOG implies that the defects in *log* are very probably caused by the decrease of active cytokinin level in the shoot meristem. However, overall cytokinin concentrations are higher in *log* mutants when the entire inflorescence is measured, presumably owing to feedback regulation (data not shown). The determination of local cytokinin contents in the tip of the shoot meristem, where *LOG* is expressed (Fig. 3), is practically impossible due to the small number of *LOG*-expressing cells. Thus, as an alternative way to examine the *in vivo* function of LOG, we examined transcript levels of the cytokinin-inducible *RESPONSE REGULATOR*[21] genes (*OsRR1* and *OsRR5*) in the meristem. The great reduction of *OsRR1* and *OsRR5* mRNAs in the floral meristem of *log-6* supports the argument that active cytokinin levels are lower in shoot meristems of *log* mutants (Fig. 4e, f). Furthermore, we asked whether there is a correlation between the LOG activity and phenotype severity in the *log* mutants (Fig. 4g). Residual activity was detected in LOG-1 and LOG-4, whose mutant alleles confer relatively weak phenotypes, but no activity was observed in LOG-5 and LOG-6, which are encoded by severe phenotype alleles (Figs 1e, 4g). These findings strongly support our notion that the *log* phenotype is the result of a malfunction of cytokinin activation in shoot meristems.

In *Arabidopsis*, two types of homeobox genes, *KNOX* and *WUSCHEL (WUS)*, function in independent and complementary pathways to establish and maintain shoot meristems[22,23]. The two pathways have direct links with cytokinins. WUS, expressed in the small number of cells within the shoot meristem, represses the type-A *RESPONSE REGULATOR* genes[24], which are the primary targets of cytokinin signal transduction[12]. KNOX proteins activate cytokinin biosynthesis in the SAM through the induction of *IPT* genes in *Arabidopsis*[25,26] and rice[27]. These findings indicate that modulation of cytokinin activity in the meristem is one of the important functions of the genetic networks controlling shoot meristem activity. Here, we have demonstrated another way in which the activity of cytokinin is regulated in the meristem through the fine-tuning of biologically active cytokinin levels. It is currently not possible to rule out the occurrence and physiological importance of a two-step reaction for cytokinin activation. In any case, the LOG-dependent direct pathway would be dominant at least in the shoot meristem. Although cytokinin functions as both a local paracrine signal and a long-distance signal[10,28], our results indicate that in the shoot meristem it functions as a local paracrine signal rather than as a long-distance signal transported from other organs. Cell-specific conversion from inactive to an active form at the site where cytokinin activity is required could be an efficient way to fully control cytokinin function, and would provide a mechanism for preventing cytokinin action in tissues or at times where it is not needed. Future analyses will determine how *LOG* function is incorporated into the overall network controlling cytokinin biosynthesis and shoot meristem activity.

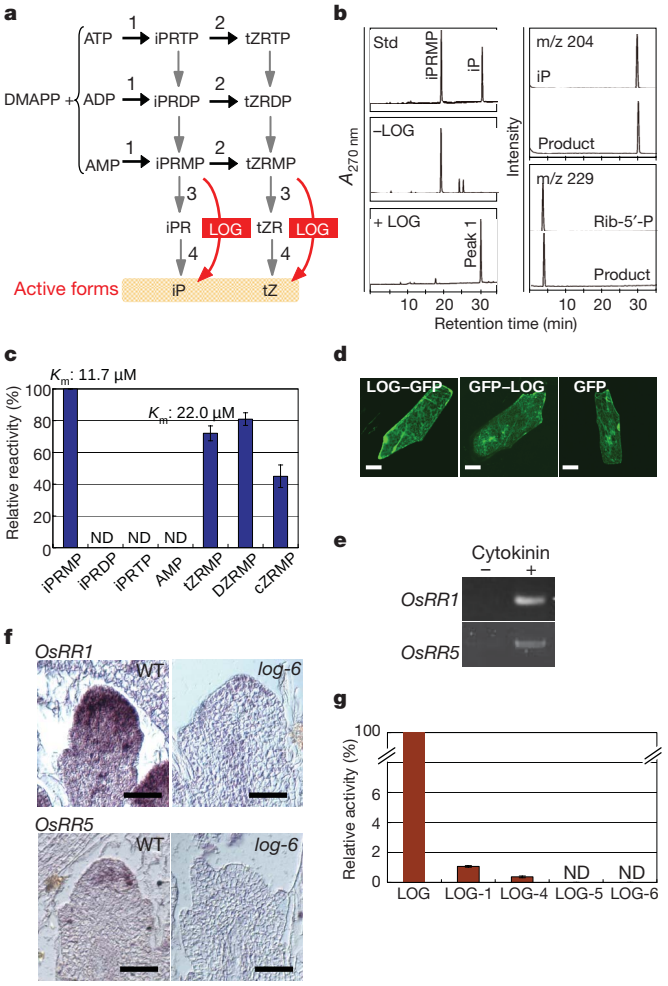


Figure 4 | Enzymatic function of LOG as a cytokinin nucleoside 5′-monophosphate phosphoribohydrolase. **a**, Schematic representation of cytokinin biosynthesis and activating pathway. DMAPP, dimethylallyl diphosphate; iPR, iP riboside; tZR, tZ riboside. 1, IPT; 2, CYP735A; 3, nucleotidase; 4, nucleosidase. Black arrows indicate reactions whose genes for the enzyme are identified, whereas grey arrows indicate that the genes are not identified. **b**, Detection and identification of reaction products. Standards (Std) of iPRMP and iP, reaction products of iPRMP without LOG (−LOG) and with LOG (+LOG) were separated by HPLC (in the left panels). Standards of iP and ribose 5′-monophosphate (Rib-5′-P) and the reaction products in +LOG (product) are analysed by liquid chromatography/mass spectrometry monitoring at mass-to-charge ratio (*m*/*z*) 204 for iP, and at 229 for Rib-5′-P (in the right panels). Mass spectra of the iP and peak 1 are shown in Supplementary Fig. 6. **c**, Substrate specificity of LOG for various nucleotides. DZRMP, dihydrozeatin riboside 5′-monophosphate; cZRMP, *cis*-zeatin riboside 5′-monophosphate. The K_m values for iPRMP and tZRMP are shown above the bars. **d**, Analysis of subcellular location of LOG protein. Chimeric constructs containing *CaMV35S::LOG–GFP* (LOG–GFP), *CaMV35S::GFP–LOG* (GFP–LOG) or *CaMV35S::GFP* (GFP) were transiently expressed in onion epidermal cells. **e**, Cytokinin responsiveness of *OsRR1* and *OsRR5* in the shoot meristem was analysed by PCR with reverse transcription (RT–PCR). RNA was isolated from shoot apices collected from plants that were treated with 10 μM benzyl aminopurine (BAP) for 60 min. **f**, *In situ* hybridization analysis of *OsRR1* and *OsRR5* expression in floral meristems of wild type (WT) and *log-6*. **g**, Enzymatic activity of LOG mutant proteins. Recombinant enzyme activity was assayed with iPRMP, and the activities are relative to LOG as 100. The data shown in **c** and **g** are means with s.d. ($n = 3$). ND, not detected. Scale bars, 50 μm (**d**); and 50 μm (**f**).

METHODS

Detailed methods are described in Supplementary Information.

Plants and mapping. *log-1* (cv. Taichung65) was generated by ethylmethane sulphonate chemical mutagenesis. *log-2*, *log-3*, and *log-6* (cv. Nipponbare) were found among a population of tissue-culture-derived plants. *log-4* and *log-5* (cv. Taichung65) were generated by *N*-methyl-*N*-nitrosourea chemical mutagenesis. The *LOG* locus was mapped using 442 mutant F_2 plants from a cross between *log-1* and wild type (Kasalath).

***In situ* hybridization.** PCR-amplified cDNAs were cloned into the pGEM-T easy vector (Promega), linearized and used to make digoxygenin-labelled anti-sense probes. Tissue fixation and *in situ* hybridization were carried out according to Kouchi *et al.*[29].

Recombinant enzymes. The coding regions of *LOG*, *log-1*, *log-4*, *log-5* or *log-6* were ligated into pCOLD1 (Takara) to express His-tagged recombinant proteins. BL21(DE3) harbouring pG-Tf2 (Takara) was used as the *Escherichia coli* host.

Enzyme assays. Enzyme activity of LOG as a cytokinin nucleoside 5′-monophosphate phosphoribohydrolase was measured by incubating the enzyme with 50 μM substrate at 30 °C for the appropriate period. The reaction products were separated and monitored as described previously[30].

Received 19 October; accepted 1 December 2006.

1. Steeves, T. A. & Sussex, I. M. *Patterns in Plant Development* (Cambridge Univ. Press, Cambridge, UK, 1989).

2. Werner, T. *et al.* Cytokinin-deficient transgenic *Arabidopsis* plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. *Plant Cell* **15,** 2532–2550 (2003).
3. Helliwell, C. A. *et al.* The *Arabidopsis AMP1* gene encodes a putative glutamate carboxypeptidase. *Plant Cell* **13,** 2115–2125 (2001).
4. Ashikari, M. *et al.* Cytokinin oxidase regulates rice grain production. *Science* **309,** 741–745 (2005).
5. Higuchi, M. *et al. In planta* functions of the *Arabidopsis* cytokinin receptor family. *Proc. Natl Acad. Sci. USA* **101,** 8821–8826 (2004).
6. Nishimura, C. *et al.* Histidine kinase homologs that act as cytokinin receptors possess overlapping functions in the regulation of shoot and root growth in *Arabidopsis*. *Plant Cell* **16,** 1365–1377 (2004).
7. Howell, S. H., Lall, S. & Che, P. Cytokinins and shoot development. *Trends Plant Sci.* **8,** 453–459 (2003).
8. Ferreira, F. J. & Kieber, J. J. Cytokinin signaling. *Curr. Opin. Plant Biol.* **8,** 518–525 (2005).
9. Mok, D. W. & Mok, M. C. Cytokinin metabolism and action. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52,** 89–118 (2001).
10. Sakakibara, H. Cytokinins: activity, biosynthesis, and translocation. *Annu. Rev. Plant Biol.* **57,** 431–449 (2006).
11. Giulini, A., Wang, J. & Jackson, D. Control of phyllotaxy by the cytokinin-inducible response regulator homologue ABPHYL1. *Nature* **430,** 1031–1034 (2004).
12. To, J. P. *et al.* Type-A *Arabidopsis* response regulators are partially redundant negative regulators of cytokinin signaling. *Plant Cell* **16,** 658–671 (2004).
13. Takei, K. *et al. AtIPT3* is a key determinant of nitrate-dependent cytokinin biosynthesis in *Arabidopsis*. *Plant Cell Physiol.* **45,** 1053–1062 (2004).
14. Miyawaki, K., Matsumoto-Kitano, M. & Kakimoto, T. Expression of cytokinin biosynthetic isopentenyltransferase genes in *Arabidopsis*: tissue specificity and regulation by auxin, cytokinin, and nitrate. *Plant J.* **37,** 128–138 (2004).
15. Werner, T., Köllmer, I., Bartrina, I., Holst, K. & Schmülling, T. New insights into the biology of cytokinin degradation. *Plant Biol. (Stuttg.)* **8,** 371–381 (2006).
16. Sentoku, N. *et al.* Regional expression of the rice *KN1*-type homeobox gene family during embryo, shoot, and flower development. *Plant Cell* **11,** 1651–1664 (1999).
17. Moriguchi, K. *et al.* The complete nucleotide sequence of a plant root-inducing (Ri) plasmid indicates its chimeric structure and evolutionary relationship between tumor-inducing (Ti) and symbiotic (Sym) plasmids in *Rhizobiaceae*. *J. Mol. Biol.* **307,** 771–784 (2001).
18. Crespi, M., Vereecke, D., Temmerman, W., Van Montagu, M. & Desomer, J. The fas operon of *Rhodococcus fascians* encodes new genes required for efficient fasciation of host plants. *J. Bacteriol.* **176,** 2492–2501 (1994).
19. Chen, C.-M. & Kristopeit, S. M. Metabolism of cytokinin: dephosphorylation of cytokinin ribonucleotide by 5′-nucleotidases from wheat germ cytosol. *Plant Physiol.* **67,** 494–498 (1981).
20. Chen, C.-M. & Kristopeit, S. M. Metabolism of cytokinin: Deribosylation of cytokinin ribonucleoside by adenosine nucleosidase from wheat germ. *Plant Physiol.* **68,** 1020–1023 (1981).
21. Jain, M., Tyagi, A. K. & Khurana, J. P. Molecular characterization and differential expression of cytokinin-responsive type-A response regulators in rice (*Oryza sativa*). *BMC Plant Biol.* **6,** 1, doi: 10.1186/1471-2229-6-1 (2006).
22. Long, J. A., Moan, E. I., Medford, J. I. & Barton, M. K. A member of the KNOTTED class of homeodomain proteins encoded by the *STM* gene of *Arabidopsis*. *Nature* **379,** 66–69 (1996).
23. Mayer, K. F. *et al.* Role of *WUSCHEL* in regulating stem cell fate in the *Arabidopsis* shoot meristem. *Cell* **95,** 805–815 (1998).
24. Leibfried, A. *et al.* WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. *Nature* **438,** 1172–1175 (2005).
25. Jasinski, S. *et al.* KNOX action in *Arabidopsis* is mediated by coordinate regulation of cytokinin and gibberellin activities. *Curr. Biol.* **15,** 1560–1565 (2005).
26. Yanai, O. *et al. Arabidopsis* KNOXI proteins activate cytokinin biosynthesis. *Curr. Biol.* **15,** 1566–1571 (2005).
27. Sakamoto, T. *et al.* Ectopic expression of KNOX homeodomain protein induces expression of cytokinin biosynthesis gene in rice. *Plant Physiol.* **142,** 54–62 (2006).
28. Faiss, M., Zalubilova, J., Strnad, M. & Schmülling, T. Conditional transgenic expression of the *ipt* gene indicates a function for cytokinins in paracrine signaling in whole tobacco plants. *Plant J.* **12,** 401–415 (1997).
29. Kouchi, H., Sekine, M. & Hata, S. Distinct classes of mitotic cyclins are differentially expressed in the soybean shoot apex during the cell cycle. *Plant Cell* **7,** 1143–1155 (1995).
30. Takei, K., Sakakibara, H. & Sugiyama, T. Identification of genes encoding adenylate isopentenyltransferase, a cytokinin biosynthesis enzyme, in *Arabidopsis thaliana*. *J. Biol. Chem.* **276,** 26405–26410 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Dennis for critical reading of this manuscript; I. Yamaguchi, T. Hashimoto and M. Nakajima for their suggestions; H. Satoh and S. Yamaki for *log* alleles; and T. Kuroha for sharing his unpublished results.

Author Contributions T.K., M.M., K.K., Y.N. and J.K. isolated, cloned and genetically characterized *LOG*, and N.U. M.K. and H.S. determined the biochemical nature of LOG. J.K. and H.S. wrote the manuscript.

Author Information Genbank accession numbers are as follows: *LOG* (AK071695), LOC_Os01g51210 (AK061091), LOC_Os05g46360 (AK068633), LOC_Os03g49050 (AC123974), LOC_Os09g37540 (AK062595), LOC_Os05g51390 (AK242659), LOC_Os03g64070 (AK061341), LOC_Os03g01880 (AK099538), LOC_Os10g33900 (AK108805), LOC_Os02g41770 (AP005000), LOC_Os04g43840 (AK069293), At2g28305 (NM_128389), At2g35990 (NM_129158), At2g37210 (NM_129277), At3g53450 (NM_115205), AT4G35190 (NM_119685), At5g03270 (NM_120405), At5g06300 (NM_120713), At5g11950 (NM_121233) and At5g26140 (NM_122515). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J. K. (akyozuka@mail.ecc.u-tokyo.ac.jp) or H.S. (sakaki@riken.jp).

LETTERS

Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas

Wen Xue[1]*, Lars Zender[1]*, Cornelius Miething[1], Ross A. Dickins[1,2], Eva Hernando[3], Valery Krizhanovsky[1], Carlos Cordon-Cardo[3] & Scott W. Lowe[1,2]

Although cancer arises from a combination of mutations in oncogenes and tumour suppressor genes, the extent to which tumour suppressor gene loss is required for maintaining established tumours is poorly understood. p53 is an important tumour suppressor that acts to restrict proliferation in response to DNA damage or deregulation of mitogenic oncogenes, by leading to the induction of various cell cycle checkpoints, apoptosis or cellular senescence[1,2]. Consequently, *p53* mutations increase cell proliferation and survival, and in some settings promote genomic instability and resistance to certain chemotherapies[3]. To determine the consequences of reactivating the p53 pathway in tumours, we used RNA interference (RNAi) to conditionally regulate endogenous p53 expression in a mosaic mouse model of liver carcinoma[4,5]. We show that even brief reactivation of endogenous p53 in p53-deficient tumours can produce complete tumour regressions. The primary response to p53 was not apoptosis, but instead involved the induction of a cellular senescence program that was associated with differentiation and the upregulation of inflammatory cytokines. This program, although producing only cell cycle arrest *in vitro*, also triggered an innate immune response that targeted the tumour cells *in vivo*, thereby contributing to tumour clearance. Our study indicates that p53 loss can be required for the maintenance of aggressive carcinomas, and illustrates how the cellular senescence program can act together with the innate immune system to potently limit tumour growth.

p53 mutations are common in human liver cancer[6], which is typically highly aggressive and resistant to non-surgical therapies. To determine the requirement for p53 loss in the maintenance of such carcinomas, we used reversible RNAi[7] to control p53 in a chimaeric liver cancer mouse model (Fig. 1a)[4,5]. Purified embryonic liver progenitor cells (hepatoblasts) were transduced with retroviruses expressing oncogenic *ras* (*HrasV12*), the tetracycline transactivator protein tTA ('tet-off') and a tet-responsive p53 miR30 design short hairpin RNA (shRNA; Supplementary Fig. 1a)[7,8], and seeded into the livers of athymic nude mice following intrasplenic injection[4,5]. To facilitate *in vivo* imaging, the oncogenic *ras* retrovirus co-expressed green fluorescent protein (GFP) and, in some experiments, hepatoblasts were also co-transduced with a luciferase reporter.

p53 expression was efficiently suppressed in the absence of doxycycline (Dox) and rapidly restored following Dox addition (Supplementary Fig. 1b, c). On transplantation into the livers of recipient mice, hepatoblast populations co-expressing Ras and the conditional p53 shRNA rapidly produced invasive hepatocarcinomas in the absence of Dox (Fig. 1b), whereas cells expressing each vector alone did not (data not shown). These tumours were GFP-positive and, if expressing luciferase, could be visualized externally by bioluminescence imaging (Fig. 1b).

Animals with advanced tumours were treated with Dox to re-establish p53 expression (Fig. 1b). Shortly after Dox administration, the p53 microRNA (miRNA) was shut off and p53 expression increased (Supplementary Fig. 1). Although tumours in untreated mice rapidly progressed, those in Dox-treated animals began to involute and became nearly undetectable within 12 days (Fig. 1b). Similar results were observed in a subcutaneous setting, where tumours could be accurately monitored using calliper measurements (Fig. 1c, left panel). Importantly, *ras*-induced liver carcinomas produced using a constitutive p53 shRNA grew similarly irrespective of Dox treatment (Fig. 1c, right), indicating that tumour regression was not due to Dox toxicity. Such regressions also occurred when p53 was reactivated in tumours co-expressing a constitutively activated *Akt* or an endogenous oncogenic *K-ras* allele and the conditional p53 shRNA (W.X., L.Z. and S.W.L., unpublished data).

To determine whether transient p53 reactivation could also cause tumour regression, we treated transformed cells in culture or tumour-bearing mice with Dox for 4 days and then removed the drug. Immunoblotting revealed that p53 could be transiently induced following Dox addition and withdrawal (Fig. 1d). In cultured cells, even two days of Dox treatment reduced colony formation to levels observed following continuous Dox treatment (Supplementary Fig. 2a). Furthermore, both *in situ* and subcutaneous liver carcinomas showed complete regressions after only four days of Dox treatment (Fig. 1e; Supplementary Fig. 2b). Thus, p53 can induce tumour involution through a process that, once activated, seems irreversible. These observations are analogous to results seen in murine tumours conditionally expressing various oncogenes, where silencing of the initiating oncogene often causes tumour regression[9,10].

The rapid involution of hepatocarcinomas re-expressing p53 is consistent with p53's well-characterized ability to promote apoptosis. We therefore examined apoptosis and proliferation in tumours before and after p53 restoration (Fig. 2a, b). Surprisingly, we observed few cells that were TUNEL-positive or contained activated caspase 3 following p53 reactivation, suggesting that the primary response to p53 was not apoptosis. Similarly, substantial necrosis was not observed in the regressing tumours. Instead, these tumours showed a marked decrease in proliferation (Ki67) that was associated with signs of cellular differentiation (Supplementary Fig. 3).

p53 can also promote cellular senescence, an apparently irreversible form of cell cycle arrest that is a potent barrier to tumorigenesis[11–14] and can be triggered by hyperactive Ras or PI3K signalling[12,15]. Interestingly, hepatocarcinomas expressing either oncogenic *ras* or *Akt* showed clear signs of senescence following p53 reactivation *in vivo* (Fig. 3a–c; data not shown for Akt), including the accumulation of senescence-associated-β-galactosidase (SA-β-Gal)

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. [2]Howard Hughes Medical Institute, Cold Spring Harbor, New York 11724, USA. [3]Division of Molecular Pathology, Memorial Sloan-Kettering Cancer Center, New York 10021, USA.
*These authors contributed equally to this work.

activity (Fig. 3a, b) and the senescence markers $p16^{INK4a}$, DcR2 and $p15^{INK4b}$ (Fig. 3c)[13,15]. SA-β-gal activity was also observed in tumours following brief Dox treatment (Fig. 3b), indicating that a pulse of p53 activity was sufficient to trigger senescence *in vivo*.

That p53 activation induces both cellular senescence and tumour involution is surprising given that senescence is a cytostatic program. Indeed, transformed cells accumulated SA-β-gal activity but subsequently remained arrested following p53 reactivation *in vitro* (Fig. 3d–f), suggesting that tumour regression involves non-cell-autonomous processes. Microscopic examination of tumours harvested at different times following p53 reactivation revealed a progressive inflammatory reaction involving polymorphonuclear leukocytes, initially developing in peri-tumoral regions and ultimately spreading throughout the tumour (Fig. 4a–f). We also observed an intense perivascular infiltration in regressing tumours, leading eventually to an overt vasculitis characterized by sclerosed vessels, hemorraghia and erythrophagocytosis (Supplementary Fig. 7). Morphological, immunofluorescence and flow cytometric analyses identified the infiltrating leukocytes as neutrophils, macrophages and natural killer cells (Supplementary Figures 4–6). These histopathological features support a model of sequential events, initiated by p53 reactivation in the tumour, activation of a dramatic inflammatory response, followed by destruction of tumour cells and neo-vasculature.

Senescent cells often acquire a gene expression signature that includes the upregulation of inflammatory cytokines and other immune modulators[16,17]. Accordingly, inflammatory cytokines known to attract macrophages (Csf1 and Mcp1), neutrophils (Cxcl1) and natural killer cells (Il15) were upregulated in liver tumours shortly following p53 reactivation (Fig. 4g). These genes were also induced by p53 in cultured hepatoma cells, demonstrating that they are expressed in tumour cells, not merely the infiltrating leukocytes (Fig. 4g). Moreover, several adhesion molecules including Icam1 (ref. 18) and Vcam1 were induced following p53 reactivation (Fig. 4g, and data not shown), indicating one way in which senescent cells could facilitate immune recognition. Finally, transcripts specific for neutrophils (Ncf2, Ncf4), macrophages (Mgl2, MSR2, CD68) and natural killer cells (Klrb1, Klrd1) were increased in senescent tumours but not cultured cells. Thus, multiple components of the innate immune system infiltrate the tumours following p53 reactivation.

To determine whether innate immune cells were required for tumour clearance, mice bearing subcutaneous hepatocarcinomas harbouring the conditional p53 shRNA were treated with gadolinium

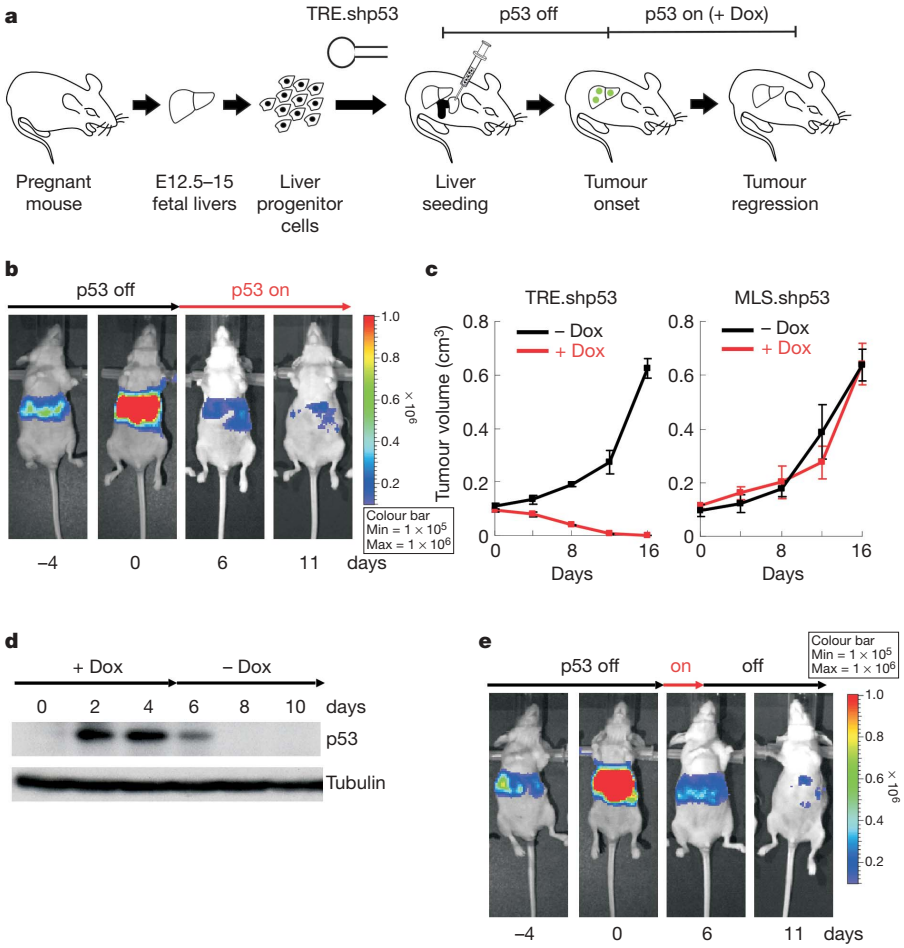


Figure 1 | Reactivation of p53 results in liver tumour regression. **a**, Embryonic liver progenitor cells were transduced with a tetracycline-regulatable p53 shRNA (TRE.shp53), tTA and H-*ras*V12. After onset of liver tumours, p53 expression could be restored by doxycycline (Dox) treatment. **b**, Reactivation of p53 leads to rapid tumour regression. Tumour-bearing mice were treated with Dox starting at day 0 and imaged at the indicated time points ($n = 9$). **c**, Subcutaneous tumours derived from *ras*-transformed liver progenitor cells with tet-off shRNA (TRE.shp53) or a non-regulatable shRNA (MLS.shp53) were grown in nude mice. Values represent mean ± s.d. ($n = 4$). **d**, p53 reactivation is reversed by Dox withdrawal. Protein lysates from liver progenitor cells pulse-treated with Dox for 4 days were immunoblotted for p53. **e**, Representative mice ($n = 6$) as in **b** were pulse-treated with Dox for 4 days and imaged at the indicated time.

chloride (a macrophage toxin)[19,20] or neutralizing antibodies to suppress neutrophil or natural killer cell function[20,21], and monitored for tumour regression following Dox treatment. All three treatments significantly delayed tumour regression following p53 reactivation (Fig. 4i), thus confirming that components of the innate immune system were actively involved in tumour clearance, presumably through a coordinated response. Importantly, each antagonist efficiently and specifically depleted the targeted immune cells from the spleen or peripheral blood (Supplementary Fig. 8), but did not prevent tumour cell senescence (Fig. 4j). Similarly, almost no tumour regression was observed following p53 reactivation in tumours grown in NOD/SCID mice —which have a highly impaired innate immune system[22]—even though p53 still induced cytostasis and senescence (Supplementary Fig. 9). Therefore, the delay in tumour clearance cannot be explained by a failure of the immune system to phagocytose dead or dying cells. Instead, these results indicate that the induction of cellular senescence and the evoked immune attack cooperate to promote tumour clearance. Of note, the athymic nude mice used here lacked functional B and T cells, which typically potentiate, but can also attenuate, inflammatory responses[23]. However, p53 reactivation in murine sarcomas also can induce senescence and tumour clearance in a completely immunocompetent setting[24].

Our study used regulatable RNAi to demonstrate that p53 loss is required for maintenance of aggressive hepatocarcinomas. We suspect that tumours harbouring *p53* mutations may be hypersensitive to restoration of p53 signalling because they have oncogenic lesions or damage signals capable of potently activating p53 (refs 3,25). Still, the consequences of restoring p53 signalling may depend on tumour origin or genotype; thus, whereas p53 reactivation induces senescence in liver carcinomas and sarcomas, lymphoid tumours respond to p53 by undergoing apoptosis[24]. Tumours may also eventually escape their dependence on *p53* mutations, but the fact that brief reactivation can cause complete tumour regressions supports the potential of transient *p53* reactivation therapies[26,27], even for advanced cancers.

Our results also identify a novel mechanism of tumour suppression involving cooperative interactions between a tumour cell senescence program and the innate immune system. They further demonstrate that, despite the cytostatic nature of the senescence program, senescent cells can turn over *in vivo*. Whether such turnover is a general feature of senescence *in vivo* is not clear[14,28], but when present may reinforce the tumour suppressive action of senescence in pre-malignant settings or in tumours following treatment with senescence- or differentiation-promoting therapies[26,29]. Conversely, our results identify a setting in which the innate immune system is provoked to coordinately attack tumour cells, presumably through both phagocytosis and direct cytotoxic killing, thereby facilitating their elimination. Although it is established that chronic inflammation triggered by senescent stromal cells or other factors can promote tumorigenesis[19,30], our study illustrates how innate immune cells—when targeted against senescent tumour cells—can have anti-tumour effects as well. Strategies that specifically harness these processes may represent a promising therapeutic approach.

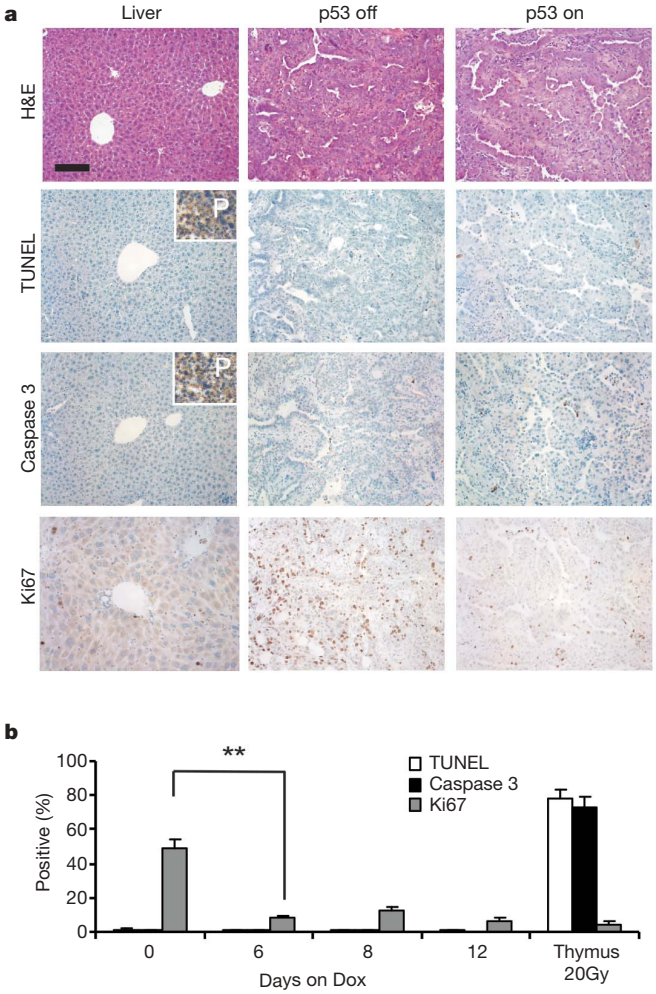


Figure 2 | The primary response to p53 reactivation is not apoptosis. **a**, Haematoxylin and eosin (H&E) immunohistochemical staining for apoptotic cells (TUNEL and Caspase 3 staining) and proliferating cells (Ki67 staining) of liver tumours before (p53 off) and after Dox treatment (p53 on). Tumours showed histopathology of human hepatocellular and cholangiocellular carcinoma. Inset (P) denotes positive controls (irradiated thymus, 20Gy (20 Gray)). 'p53 on' shows representative tumour on day 6. Scale bar, 100 μm. **b**, Quantification of **a**. Values represent mean ± s.d. ($n = 4$; $^{**}P < 0.002$).

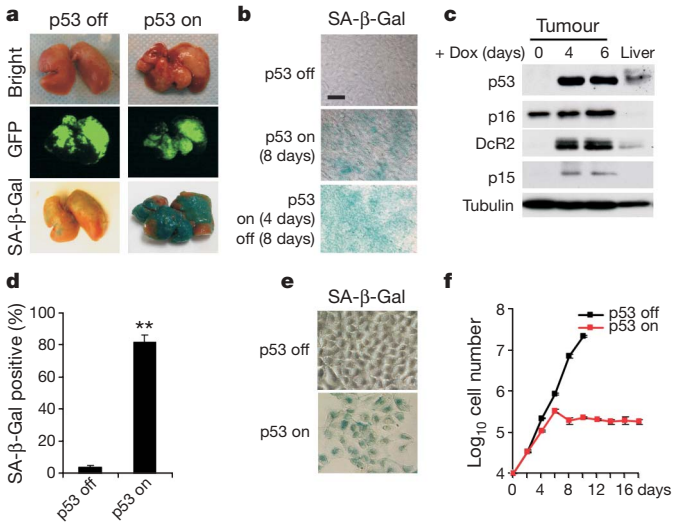


Figure 3 | p53 reactivation induces cellular senescence. **a**, SA-β-Gal staining of representative tumour-bearing livers untreated (p53 off) or treated with Dox (p53 on, day 6) ($n = 3$). **b**, SA-β-Gal staining of tumour sections ($n = 3$). Tumours were either untreated (p53 off), constantly treated with Dox for 8 days (p53 on 8 days) or briefly treated for 4 days and left untreated for 8 days (p53 on 4 days/off 8 days). Scale bar, 25 μm. **c**, Immunoblotting for senescence markers in normal liver or liver tumours treated with Dox for 0, 4 and 6 days. **d**, Liver progenitor cells harbouring *ras* and tet-off shp53 were cultured in Dox-containing medium for 6 days (p53 on) and stained for SA-β-Gal. Values represent mean ± s.d. ($n = 3$; $^{**}P < 0.0001$). **e**, Representative pictures from **d**. **f**, Cells as in **d** were cultured with (p53 on, red line) or without Dox (p53 off, black line) and cell numbers were counted. Values represent mean ± s.d. ($n = 4$).

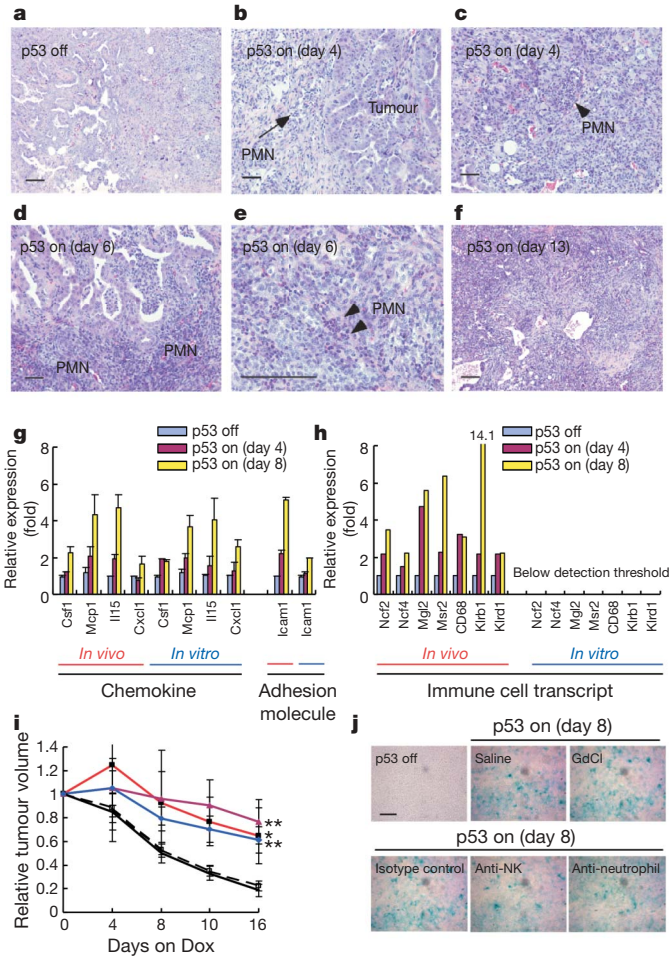


Figure 4 | Clearance of liver tumours by an innate immune response. **a–f**, Progressive immune infiltration in regressing tumours following p53 reactivation (H&E; $n = 4$). Arrows denote peri-tumoral polymorphonuclear (PMN) leukocytes. Arrowheads denote intra-tumoral polymorphonuclear leukocytes. Scale bar, 100 μm. **e**, High magnification view of **d**. **g**, p53 reactivation leads to increased expression of chemokines and adhesion molecules in senescent tumours (*in vivo*) and cultured liver progenitor cells (*in vitro*). Values represent mean ± s.d. (duplicate samples with triplicate qPCR). All the D4 and D8 values are statistically significant compared with D0 ($P < 0.05$). **h**, p53 reactivation is accompanied by increased immune cell transcripts in senescent tumours (*in vivo*) but not in cultured cells (*in vitro*). Values represent the average of duplicate samples from microarrays. **i**, Subcutaneous hepatocarcinomas were treated with Dox to induce tumour regression. The macrophage toxin GdCl (red line), an anti-neutrophil antibody (purple) or an anti-natural killer-cell antibody (blue) was administered. Saline (solid black line) or an isotype control antibody (dashed black line) served as controls. Values represent mean ± s.d. ($n = 4$; $^*P < 0.02$; $^{**}P < 0.002$ for day 16). **j**, Blockade of innate immune cells does not prevent p53-induced senescence. Frozen sections from control animals or immune antagonist treated animals were stained for SA-β-Gal activity ($n = 3$). Scale bar, 50 μm.

METHODS

See Supplementary Methods for detailed experimental methods.

Generation and analysis of liver tumours. Isolation, culture and retroviral infection of murine hepatoblasts were as described[4,5]. Doxycycline treatment (BD Biosciences) and bioluminescence imaging (Xenogen) was performed according to the manufacturer's instructions. Histopathological evaluation was performed by an experienced pathologist (C.C.C.). Immunohistochemistry was performed as described[5].

Tumour characterization. Fresh tumour tissue was lysed in Laemmli buffer using a tissue homogenizer and analysed by immunoblotting by standard procedures with the indicated antibodies. Detection of SA-β-gal activity was performed as described at pH 5.5 (ref. 7). Sections (10 μm) of snap frozen tumour tissue were fixed with 1% formalin for 1 min and stained for 12 h. Tumour-bearing livers were fixed with 4% formalin overnight and stained for 4 h. Cultured cells were fixed with 4% formalin for 5 min and stained for 10 h.

RNA expression analyses. RNA isolation (Qiagen) and TaqMan reverse transcriptase reaction (Applied Bosystems) were according to the manufacturer's instructions. Quantitative PCR (qPCR) reactions (Bio-Rad) for each sample were done in triplicate. Microarray experiments were performed on Mouse Genome 430A 2.0 arrays (Affymetrix).

Data analysis. All the statistical analysis was done by Student's *t*-test.

Received 26 September; accepted 13 December 2006.
Published online 24 January 2007.

1. Harris, S. L. & Levine, A. J. The p53 pathway: positive and negative feedback loops. *Oncogene* **24**, 2899–2908 (2005).
2. Sherr, C. J. Principles of tumor suppression. *Cell* **116**, 235–246 (2004).
3. Lowe, S. W., Cepero, E. & Evan, G. Intrinsic tumour suppression. *Nature* **432**, 307–315 (2004).
4. Zender, L. *et al.* Generation and analysis of genetically defined liver carcinomas derived from bipotential liver progenitors. *Cold Spring Harb. Symp. Quant. Biol.* **70**, 251–261 (2005).
5. Zender, L. *et al.* Identification and validation of oncogenes in liver cancer using an integrative oncogenomic approach. *Cell* **125**, 1253–1267 (2006).
6. Staib, F., Hussain, S. P., Hofseth, L. J., Wang, X. W. & Harris, C. C. TP53 and liver carcinogenesis. *Hum. Mutat.* **21**, 201–216 (2003).
7. Dickins, R. A. *et al.* Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nature Genet.* **37**, 1289–1295 (2005).
8. Silva, J. M. *et al.* Second-generation shRNA libraries covering the mouse and human genomes. *Nature Genet.* **37**, 1281–1288 (2005).
9. Chin, L. *et al.* Essential role for oncogenic Ras in tumour maintenance. *Nature* **400**, 468–472 (1999).
10. Jain, M. *et al.* Sustained loss of a neoplastic phenotype by brief inactivation of MYC. *Science* **297**, 102–104 (2002).
11. Braig, M. *et al.* Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature* **436**, 660–665 (2005).
12. Chen, Z. *et al.* Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. *Nature* **436**, 725–730 (2005).
13. Collado, M. *et al.* Tumour biology: senescence in premalignant tumours. *Nature* **436**, 642 (2005).
14. Michaloglou, C. *et al.* $BRAF^{E600}$-associated senescence-like cell cycle arrest of human naevi. *Nature* **436**, 720–724 (2005).
15. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic *ras* provokes premature cell senescence associated with accumulation of p53 and $p16^{INK4a}$. *Cell* **88**, 593–602 (1997).
16. Minamino, T. *et al.* Ras induces vascular smooth muscle cell senescence and inflammation in human atherosclerosis. *Circulation* **108**, 2264–2269 (2003).
17. Shelton, D. N., Chang, E., Whittier, P. S., Choi, D. & Funk, W. D. Microarray analysis of replicative senescence. *Curr. Biol.* **9**, 939–945 (1999).
18. Gorgoulis, V. G. *et al.* p53-dependent ICAM-1 overexpression in senescent human cells identified in atherosclerotic lesions. *Lab. Invest.* **85**, 502–511 (2005).
19. Maeda, S., Kamata, H., Luo, J. L., Leffert, H. & Karin, M. IKKβ couples hepatocyte death to cytokine-driven compensatory proliferation that promotes chemical hepatocarcinogenesis. *Cell* **121**, 977–990 (2005).
20. Mundt, B. *et al.* Involvement of TRAIL and its receptors in viral hepatitis. *FASEB J.* **17**, 94–96 (2003).
21. Hong, F. *et al.* β-glucan functions as an adjuvant for monoclonal antibody immunotherapy by recruiting tumoricidal granulocytes as killer cells. *Cancer Res.* **63**, 9023–9031 (2003).
22. Shultz, L. D. *et al.* Multiple defects in innate and adaptive immunologic function in NOD/LtSz-scid mice. *J. Immunol.* **154**, 180–191 (1995).
23. Ghiringhelli, F., Menard, C., Martin, F. & Zitvogel, L. The role of regulatory T cells in the control of natural killer cells: relevance during tumor progression. *Immunol. Rev.* **214**, 229–238 (2006).
24. Ventura, A. *et al.* Restoration of p53 function leads to tumour regression *in vivo*. *Nature* advance online publication, doi:10.1038/nature05541 (24 January 2007).
25. Christophorou, M. A., Ringshausen, I., Finch, A. J., Swigart, L. B. & Evan, G. I. The pathological response to DNA damage does not contribute to p53-mediated tumour suppression. *Nature* **443**, 214–217 (2006).
26. Bykov, V. J. *et al.* Restoration of the tumor suppressor function to mutant p53 by a low-molecular-weight compound. *Nature Med.* **8**, 282–288 (2002).
27. Vassilev, L. T. *et al.* *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844–848 (2004).
28. Gray-Schopfer, V. C. *et al.* Cellular senescence in naevi and immortalisation in melanoma: a role for p16? *Br. J. Cancer* **95**, 496–505 (2006).
29. Roninson, I. B. Tumor cell senescence in cancer treatment. *Cancer Res.* **63**, 2705–2715 (2003).

30. Krtolica, A., Parrinello, S., Lockett, S., Desprez, P. Y. & Campisi, J. Senescent fibroblasts promote epithelial cell growth and tumorigenesis: a link between cancer and aging. *Proc. Natl Acad. Sci. USA* **98,** 12072–12077 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Bianco and M. Jiao for technical assistance. We also thank G. Evan, T. Jacks, A. Ventura, M. Narita, A. Chicas, M. Yon, G. Hannon and other members of the Lowe and Hannon laboratories for advice and discussions. We thank M. McCurrach for editorial assistance. W.X. is in the MCB graduate program at Stony Brook University. This work was generously supported by the Emmy Noether Programme of the German Research Foundation, Alan and Edith Seligson, the Don Monti Foundation, and grants from the National Institutes of Health (C.C.C, S.W.L.). This work is dedicated to our friend and colleague Dr. Enrique (Henry) Cepero.

Author Contributions W.X.: study design and conduction of experiments; L.Z.: study design and conduction of experiments; C.M.: design and conduction of flow cytometry experiments; R.A.D.: vector development; E.H.: histopathological analyses; V.K.: microarray analysis; C.C.C: histopathological analyses; S.W.L.: study design, principal investigator.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.W.L. (lowe@cshl.edu).

LETTERS

Restoration of p53 function leads to tumour regression *in vivo*

Andrea Ventura[1]*, David G. Kirsch[1,2]*, Margaret E. McLaughlin[1], David A. Tuveson[1], Jan Grimm[3], Laura Lintault[1], Jamie Newman[1], Elizabeth E. Reczek[1], Ralph Weissleder[3] & Tyler Jacks[1,4]

Tumorigenesis is a multi-step process that requires activation of oncogenes and inactivation of tumour suppressor genes[1]. Mouse models of human cancers have recently demonstrated that continuous expression of a dominantly acting oncogene (for example, *Hras*, *Kras* and *Myc*) is often required for tumour maintenance[2–5]; this phenotype is referred to as oncogene addiction[6]. This concept has received clinical validation by the development of active anticancer drugs that specifically inhibit the function of oncoproteins such as BCR-ABL, c-KIT and EGFR[7–10]. Identifying additional gene mutations that are required for tumour maintenance may therefore yield clinically useful targets for new cancer therapies. Although loss of p53 function is a common feature of human cancers[11], it is not known whether sustained inactivation of this or other tumour suppressor pathways is required for tumour maintenance. To explore this issue, we developed a Cre-*loxP*-based strategy to temporally control tumour suppressor gene expression *in vivo*. Here we show that restoring endogenous p53 expression leads to regression of autochthonous lymphomas and sarcomas in mice without affecting normal tissues. The mechanism responsible for tumour regression is dependent on the tumour type, with the main consequence of p53 restoration being apoptosis in lymphomas and suppression of cell growth with features of cellular senescence in sarcomas. These results support efforts to treat human cancers by way of pharmacological reactivation of p53.

Biochemical and genetic studies have demonstrated that p53 responds to genotoxic and oncogenic stresses by inducing cell cycle arrest or apoptosis[11]. Because oncogenic stress can persist after p53 inactivation, loss of p53 function may not only play a role in the early stages of tumour development, but also be required for the continued proliferation or survival of an established tumour[12–14]. To test this hypothesis, we used a genetic strategy to restore endogenous p53 expression in primary, autochthonous tumours. Mice carrying a reactivatable p53 knockout allele were generated by inserting a transcription–translation stop cassette flanked by *loxP* sites (LSL) in the first intron of the endogenous wild-type p53 locus (Fig. 1a, b). When in place, the STOP cassette efficiently prevents expression of the *p53* gene (Fig. 1c, d). Cells from homozygous $p53^{LSL/LSL}$ (referred to hereafter as p53-LSL) mice are functionally equivalent to p53 null ($p53^{-/-}$) cells, as demonstrated by the absence of cell-culture-induced senescence (Fig. 1e) and by their marked genetic instability

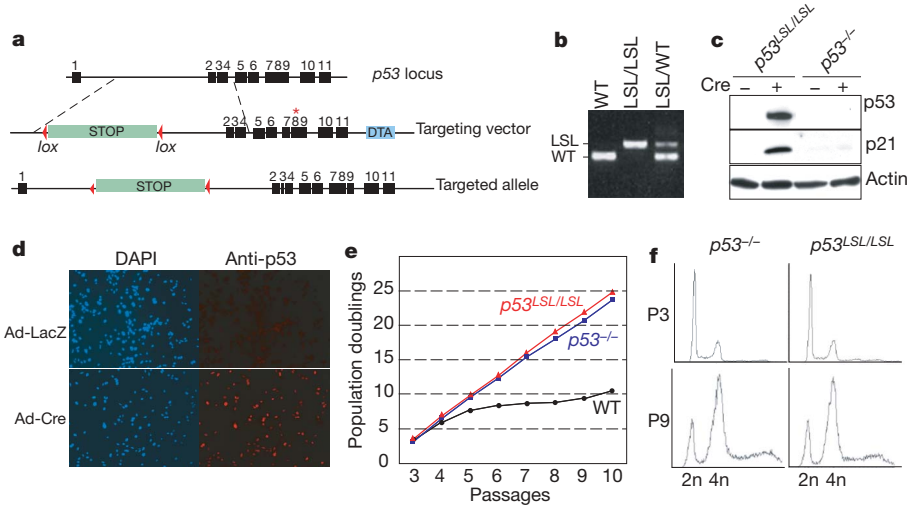


Figure 1 | Generation of the p53-LSL allele. a, The p53-LSL allele was generated while engineering a point-mutant p53 allele[31] when homologous recombination occurred between the LSL cassette and the point mutation (asterisk). **b**, Genotyping of p53-LSL mice by PCR (polymerase chain reaction) amplification of tail DNA. **c**, Western blot of passage 5 MEFs infected with empty or Cre-expressing recombinant adenoviruses (Adeno-empty and Adeno-Cre). **d**, Anti-p53 immunostaining of $p53^{LSL/LSL}$ (that is, p53-LSL) MEFs infected with Adeno-LacZ or Adeno-Cre and then treated with doxorubicin (0.2 μg ml^{-1}; 13 h). DAPI, 4,6-diamidino-2-phenylindole. **e**, 3T3 protocol on $p53^{LSL/LSL}$, $p53^{-/-}$ and wild-type (WT) MEFs. **f**, DNA content of early (P3) and late (P9) passage $p53^{LSL/LSL}$ or $p53^{-/-}$ MEFs. Diploid (2n) and tetraploid (4n) DNA content is indicated.

[1]Center for Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA. [2]Department of Radiation Oncology, [3]Center for Molecular Imaging Research, Massachusetts General Hospital, Boston, Massachusetts 02129, USA, and Harvard Medical School, Boston, Massachusetts 02115, USA. [4]Howard Hughes Medical Institute, Chevy Chase, Maryland 20815, USA.
*These authors contributed equally to this work.

(Fig. 1f). However, owing to the presence of flanking *loxP* sites, the STOP cassette can be excised by the Cre recombinase[15], thus restoring expression of the endogenous *p53* gene (Fig. 1c, d). Because mice lacking p53 are tumour-prone[16,17], this system offers an opportunity to study the consequences of p53 reactivation in primary tumours *in vivo*.

To temporally control p53 reactivation *in vivo*, we have also generated mice carrying a Cre-recombinase-Oestrogen-Receptor-T2 (Cre-ER^{T2}) allele targeted to the ubiquitously expressed ROSA26 locus (Supplementary Fig. 1a, b). The ER^{T2} moiety retains the Cre recombinase in the cytoplasm until tamoxifen administration releases this inhibition[18], thus permitting the recombination of genomic *loxP* sites. Efficient tamoxifen-induced Cre-mediated recombination throughout the body was demonstrated by crossing Cre-ER^{T2} mice to mice carrying a Cre-responsive β-galactosidase reporter allele[19] (Supplementary Fig. 1c).

Next, Cre-ER^{T2} and p53-LSL mice were crossed to generate cohorts of p53-LSL homozygous mutant animals carrying the Cre-ER^{T2} allele (*p53*$^{LSL/LSL}$;Cre-recombinase-Oestrogen-Receptor-T2, hereafter referred to as p53-LSL;Cre-ER^{T2}) as well as p53-LSL homozygotes without the Cre-ER^{T2} allele (hereafter referred to as p53-LSL). To accelerate tumour formation, some of these mice were irradiated shortly after birth, as radiation decreases tumour latency in p53 null mice[20]. These animals underwent periodic magnetic resonance imaging (MRI) to detect the presence of cancer (Fig. 2a). After tumour detection, the mice were treated with tamoxifen and then re-imaged at different times thereafter.

As shown in Fig. 2, in the majority of p53-LSL;Cre-ER^{T2} mice, tamoxifen treatment caused regression of autochthonous lymphomas and sarcomas (Fig. 2b, c, f; Supplementary Fig. 2, Supplementary Movies). This effect was due to p53 reactivation rather than tamoxifen treatment alone, because in Cre-ER^{T2}-negative p53-LSL mice, tumours rapidly progressed despite treatment (Fig. 2d–f). Figure 2f summarizes the responses observed in the animals studied. Out of 10 Cre-ER^{T2}-positive tumours treated (6 thymic lymphomas, 3 sarcomas and 1 intra-abdominal lymphoma), 7 showed regression ranging from 46% to 100%. In two tumours (a thymic lymphoma and an osteosarcoma), tamoxifen prevented tumour progression but failed to decrease the size of the neoplasia (Fig. 2f; Supplementary Fig. 3). In one thymic lymphoma tumour progression was observed despite tamoxifen treatment (Fig. 2f; Supplementary Fig. 3). DNA and immunohistochemical analysis indicated that in this case resistance to tamoxifen was due to the specific deletion of the Cre-ER^{T2} allele in the tumour (Supplementary Fig. 3e, f). In contrast, all Cre-ER^{T2}-negative tumours (two thymic lymphomas, two abdominal lymphomas and three sarcomas) progressed despite tamoxifen administration (Fig. 2d–f). These results demonstrate that sustained p53 inactivation is required for tumour maintenance in autochthonous lymphomas and sarcomas in the mouse.

To study the mechanisms responsible for tumour regression, tumours from p53-LSL;Cre-ER^{T2} mice were analysed 24 or 48 h after tamoxifen treatment. In thymic lymphomas, tumour regression was already apparent at 48 h (approximately 75% reduction in tumour volume; see Fig. 3a). In addition, p53 expression was detectable as early as 24 h following tamoxifen treatment and was associated with widespread apoptosis (Fig. 3b). We next derived three thymic lymphoma cell lines, two from *p53*$^{LSL/+}$;Cre-ER^{T2} mice and one from a p53-LSL;Cre-ER^{T2} mouse. An additional line, to be used as a control, was derived from a thymic lymphoma arising in a *p53*$^{+/-}$ mouse. In all cases, p53 expression was undetectable by western blot (data not shown), suggesting that in the tumours arising in *p53*$^{LSL/+}$ and *p53*$^{+/-}$ mice expression of the wild-type p53 allele had been lost. 4-hydroxytamoxifen (4-OHT) treatment caused cell death in the p53-LSL;Cre-ER^{T2}-positive cell lines but not in the control cells (Fig. 3c). A time-course experiment with the p53-LSL;Cre-ER^{T2} thymic lymphoma cell line showed that cell death began within 12 h after 4-OHT administration and was virtually complete by 96 h (Fig. 3d). Moreover, just as the primary thymic lymphomas showed apoptotic

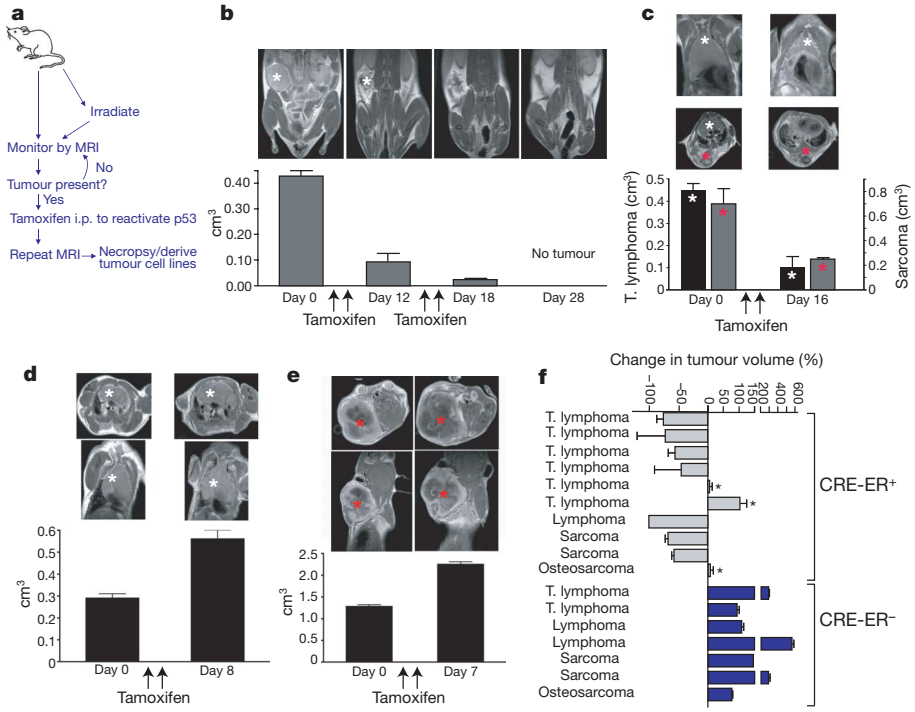


Figure 2 | p53 restoration leads to tumour regression *in vivo*. **a**, Flow chart of the strategy used to determine tumour response. i.p., intraperitoneal. **b–d**, MRI images (top) and tumour volumes (bottom) of p53-LSL;Cre-ER^{T2} (**b**, **c**) and p53-LSL (**d**, **e**) mice in response to tamoxifen (arrows). The tumours (asterisks) were an abdominal lymphoma (**b**), two thymic lymphomas (t. lymphoma; **c** and **d**, white asterisks) and two sarcomas (**c** and **e**, red asterisks). The volumes were calculated from the available MRI sequences ($n = 2$ to 6) for each time point, and are shown as mean + 1 s.d. **f**, Summary of maximal responses to tamoxifen of tumours from Cre-ER^{T2}-positive (grey bars) and Cre-ER^{T2}-negative (blue bars) mice. Asterisks indicate tumours from Cre-ER^{T2}-positive mice with limited or no response (see also Supplementary Fig. S3).

cell death (Fig. 3b), immunoblot for activated caspase-3 and flow-cytometry analysis of annexin-5 positive cells (Fig. 3e and data not shown) demonstrated that apoptosis was the primary consequence of p53 reactivation in this context. Interestingly, the cell cycle distribution of the viable lymphoma cells did not change in response to 4-OHT administration (Fig. 3e), despite induction of the p21 cyclin-dependent kinase inhibitor (data not shown).

Compared to lymphomas, sarcoma regression upon tamoxifen treatment was more delayed, with only a modest reduction at 48 h (Fig. 4a). Furthermore, although p53 was clearly expressed in the sarcomas at 24 and 48 h (Fig. 4b and data not shown) extensive apoptosis was not evident at these time points (Fig. 4b). Despite the absence of a clear increase in apoptosis, prolonged tamoxifen treatment led to regression of sarcomas (Fig. 2c, f; Supplementary Figs 4 and 5), with the residual mass largely composed of necrotic tissue (Supplementary Fig. 5). In order to determine the mechanism underlying sarcoma regression, we investigated the consequences of p53 restoration in two cell lines derived from p53-LSL sarcomas.

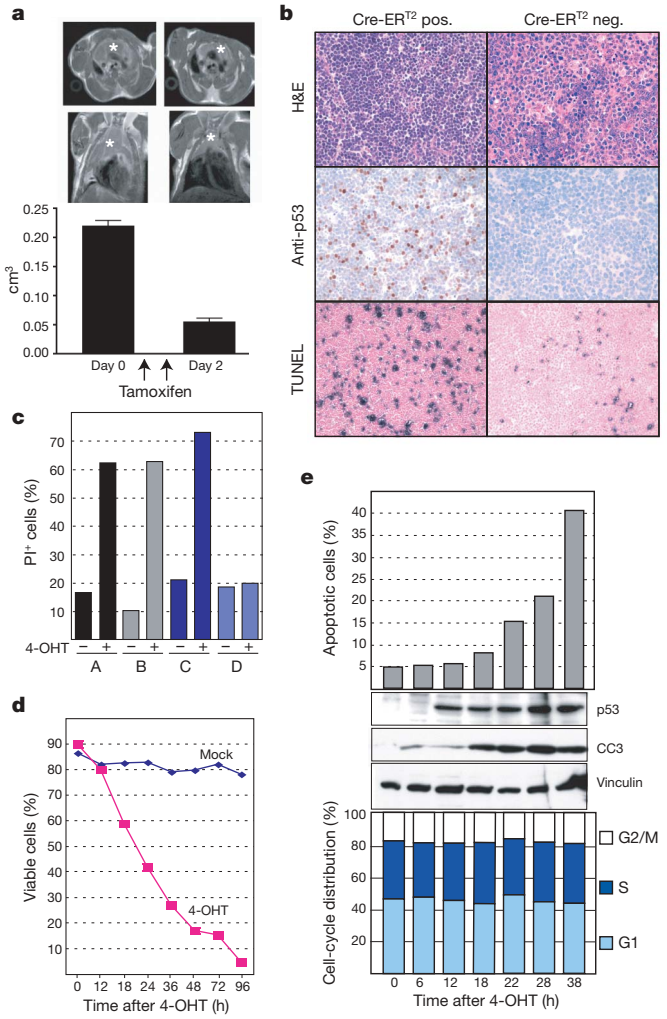


Figure 3 | p53 restoration in lymphomas leads to apoptosis. **a**, MRI images (top) and tumour volumes (bottom) of a p53-LSL;Cre-ER^{T2} lymphoma. Error bars, + 1 s.d. **b**, p53 immunohistochemistry and TUNEL staining in thymic lymphomas after tamoxifen treatment. H&E, haematoxylin and eosin. **c**, Viability of thymic lymphoma cell lines derived from Cre-ER^{T2}-positive mice (lines A–C), and a *p53*$^{+/-}$ mouse (line D) treated with vehicle (ethanol) or 250 nM 4-OHT for 48 h. Dead cells were identified by their retention of propidium iodide (PI+). **d**, Time course of cell viability of line A treated with vehicle (mock) or 250 nM 4-OHT. **e**, Line A was treated with 4-OHT and analysed for apoptosis (annexin V staining, top panel), expression of p53, cleaved caspase 3 (CCR) and vinculin (western blots, middle panel) and cell cycle distribution (PI staining, bottom panel).

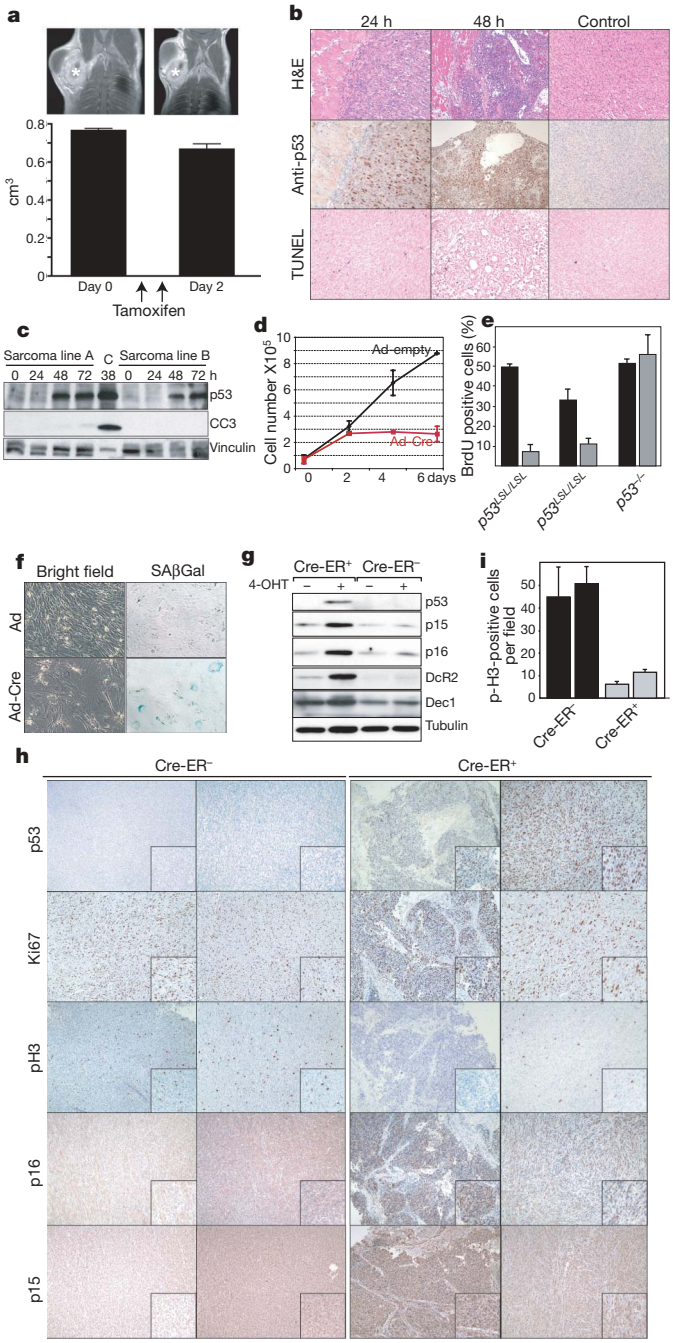


Figure 4 | p53 restoration in sarcomas leads to growth suppression with features of cellular senescence. **a**, MRI images (top) and tumour volumes (bottom) of a p53-LSL;Cre-ER^{T2} soft tissue sarcoma (STS). Error bars, + 1 s.d. **b**, p53 immunohistochemistry and TUNEL staining of p53-LSL;Cre-ER^{T2} STSs after tamoxifen treatment. The control is a Cre-ER^{T2}-negative sarcoma. **c**, p53-LSL sarcoma cells were infected with Adeno-Cre and harvested for immunoblot. Lane C, positive control for cleaved caspase 3. **d**, Proliferation assay. Average of two independent cell lines ± 1 s.d. is shown. **e**, BrdU (5-bromodeoxyuridine) incorporation in sarcoma cells infected with Adeno-GFP (black bars) or Adeno-Cre-GFP (grey bars). Error bars, + 1 s.d. **f**, Senescence-associated β-galactosidase (SAβGal) staining of a sarcoma cell line 5 days after infection with Adeno-empty or Adeno-Cre. **g**, Immunoblot analysis of senescence markers in p53-LSL;Cre-ER^{T2} and p53-LSL sarcoma cells 11 days after 4-OHT tamoxifen treatment. **h**, Immunohistochemistry analysis of four sarcomas 5–6 days after *in vivo* tamoxifen administration. **i**, Quantification of phospho-histone H3 positive cells in sarcomas from **h**. Average of seven random 40× fields + 1 s.d.

Consistent with the results *in vivo*, expression of Cre recombinase in these cell lines failed to induce significant apoptosis (Fig. 4c). Instead, restoration of p53 expression suppressed proliferation (Fig. 4d) and induced cell cycle arrest (Fig. 4e). In addition, these cells lost their spindle morphology, appeared flat and enlarged, and many of them expressed senescence-associated β-galactosidase (Fig. 4f).

Recent work has led to the identification of a number of proteins whose expression is increased in senescent pre-neoplastic lesions[21]. These markers include the cdk-inhibitors p15-Ink4b and p16-Ink4a, as well as DcR2 and Dec1. We examined their expression in sarcoma cell lines derived from p53-LSL;Cre-ER^{T2} and p53-LSL mice. Only the Cre-ER^{T2}-positive sarcoma cells induced p15-Ink4b, p16-Ink4a, DcR2 and (to a lesser extent) Dec-1 in response to tamoxifen treatment (Fig. 4g). We next examined whether cell cycle arrest and senescence markers are similarly induced by p53 restoration in sarcomas *in vivo*. Senescence-associated β-galactosidase positive cells were observed in one out of three Cre-ER^{T2}-positive sarcomas treated with tamoxifen for 5–6 days, but in none of the Cre-ER^{T2}-negative controls (Supplementary Fig. S6, and data not shown). Consistent with the data obtained in sarcoma cell lines, p15-Ink4b and p16-Ink4a expression was higher in the tamoxifen-treated Cre-ER^{T2}-positive sarcomas, compared to the Cre-ER^{T2}-negative ones (Fig. 4h). Although Ki67 staining was not significantly different after p53 restoration in these tumours, an analysis of phospho-histone H3 revealed a markedly lower fraction of proliferation in the tamoxifen-treated Cre-ER^{T2}-positive sarcomas (Fig. 4h and i). Taken together, these results indicate that the main consequence of p53 restoration in sarcomas is cell cycle arrest with some features of senescence. We did not observe upregulation of all senescence markers in every sarcoma studied *in vivo*, which may reflect the genetic and histological heterogeneity between the spontaneous mesenchymal tumours analysed here and epithelial pre-neoplastic lesions reported by others[21]. Furthermore, owing to the relatively low incidence of sarcomas in p53 null animals, our analysis of senescence markers *in vivo* has been relatively limited to date. We anticipate that with further studies of p53 restoration in sarcomas, we will be able to clarify the arrest/senescent phenotype of different sarcoma subtypes more fully.

Although tamoxifen treatment of p53-LSL;Cre-ER^{T2} mice leads to excision of the STOP cassette in both neoplastic and normal tissues, expression of the p53 protein was detectable by immunohistochemistry only in tumour cells (Figs 3b and 4b, and data not shown). Furthermore, histopathological analysis revealed no signs of p53-mediated toxicity in normal organs after acute p53 reactivation (data not shown). These results indicate that in p53 null mice, transformed cells are uniquely primed to stabilize and activate p53. In established tumours, one potential stimulus for p53 protein stabilization and activation is persistent oncogenic stress. Because the tumour suppressor gene *p19^{Arf}* is induced in response to oncogenic stimuli and is a well known activator of p53 (ref. 22), we analysed its expression levels in tumours from p53-LSL animals. Strikingly, high p19^{Arf} expression was frequently observed in both lymphomas and sarcomas but not in the normal tissues analysed (Supplementary Fig. 7 and data not shown). These results suggest that in lymphomas and sarcomas, p19^{Arf} might prime tumour cells to respond to p53 reactivation.

In this study, we have used a novel genetic strategy for restoring endogenous p53 activity in cells and in mice. By combining a reactivatable loss-of-function p53 allele together with a temporally regulated Cre recombinase, we demonstrate that sustained inactivation of p53 is required for the maintenance of primary, autochthonous lymphomas and sarcomas. This approach may be of general application for the study of other tumour suppressor gene pathways.

Interestingly, the mechanism of tumour regression after p53 restoration appears to be tumour type specific. In lymphomas, restoration of p53 caused widespread apoptosis, whereas the major consequence of p53 reactivation in sarcomas was cell cycle arrest with features of cellular senescence. Because sarcomas also regressed following p53 reactivation *in vivo*, it is possible that senescent cells are rapidly cleared from the tumour mass. This hypothesis is consistent with results obtained in a mouse model of liver cancer in an accompanying paper[23]. Alternatively, sarcoma regression may be a consequence of p53-dependent effects on tumour vasculature[24,25] or on other stromal components. The effect is also tumour specific in that normal cells appeared not to respond to p53 reactivation. Thus, this represents an extremely promising anticancer strategy with a broad therapeutic window.

These results are also consistent with work in other model systems. For example, mouse embryo fibroblasts (MEFs) transformed by ectopically expressing an oncogenic *K-Ras* allele in the setting of doxycycline-regulated p53 knock-down enter senescence upon restoration of p53 expression[26]. Tumours obtained by injecting these cells into nude mice also regressed in response to doxycycline treatment[26]. Likewise, immortalized MEFs from a p53–oestrogen-receptor(tamoxifen) fusion knock-in mouse strain undergo senescence after treatment with tamoxifen[27], and temporarily activating p53 in these mice delays tumour development[28].

Owing to the prevalence of p53 pathway inactivation in human cancers, several pharmacological strategies aimed at restoring p53 function have been proposed. These include small molecules that restore point-mutant p53 proteins to a transcriptionally competent conformation[13], as well as compounds that interfere with the Mdm2–p53 interaction[29] and gene-therapy-based approaches aimed at introducing a wild-type copy of the p53 gene into tumour cells[30]. Provided that human cancers, like the mouse cancers studied here, remain dependent on sustained p53 inactivation for tumour maintenance, our results lend strong support to such therapeutic efforts.

METHODS

A detailed description of materials and methods is given in Supplementary Information.

Generation of p53-LSL mice. The p53-LSL allele was obtained as a by-product of our efforts to generate a R270H point mutant p53 allele[31]. All p53 exons and intron–exon boundaries were sequenced to ensure the absence of mutations. Oligonucleotide sequences for genotyping are available on request.

Generation of R26-Cre-ER^{T2} mice. A plasmid containing the Cre-ER^{T2} complementary DNA (pCre-ER^{T2}) was obtained from the laboratory of P. Chambon[19]. This cDNA was targeted to the ROSA26 locus as described in detail in the Supplementary Information.

Tamoxifen treatment. All animal studies and procedures were approved by the MIT Institutional Animal Care and Use Committee, and by the Subcommittee on Research Animal Care at Massachusetts General Hospital. Mice were of a mixed 129Sv/Jae and C57/B6 background. After detecting a tumour by MRI, mice were treated with tamoxifen (Sigma) by intraperitoneal injection. 100 μl of tamoxifen (10 mg ml^{-1} in corn oil) was injected every two to three days. Follow-up MRI was generally obtained 7–10 days after the first tamoxifen treatment.

MRI and tumour volume analysis. All animals were sequentially imaged using a 4.7 T Bruker Pharmascan (Bruker BioSpin) to screen for tumour growth and treatment response to tamoxifen. Tumour volume measurements were performed using T1- and T2-weighted coronal and axial image stacks. The tumours were manually segmented using Amira software (TGS) to obtain the tumour volume in cm^3. For each time point, all measurements from the available sequences were used to calculate a mean volume ± standard deviation in order to limit errors due to the manual segmentation process.

Received 26 September; accepted 13 December 2006.
Published online 24 January 2007.

1. Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *Cell* **100**, 57–70 (2000).
2. Chin, L. *et al.* Essential role for oncogenic Ras in tumour maintenance. *Nature* **400**, 468–472 (1999).
3. Jain, M. *et al.* Sustained loss of a neoplastic phenotype by brief inactivation of MYC. *Science* **297**, 102–104 (2002).
4. Fisher, G. H. *et al.* Induction and apoptotic regression of lung adenocarcinomas by regulation of a *K-Ras* transgene in the presence and absence of tumor suppressor genes. *Genes Dev.* **15**, 3249–3262 (2001).
5. Pelengaris, S., Khan, M. & Evan, G. I. Suppression of Myc-induced apoptosis in β cells exposes multiple oncogenic properties of Myc and triggers carcinogenic progression. *Cell* **109**, 321–334 (2002).

6. Weinstein, I. B. Cancer. Addiction to oncogenes–the Achilles heel of cancer. *Science* **297,** 63–64 (2002).
7. Druker, B. J. *et al.* Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N. Engl. J. Med.* **344,** 1031–1037 (2001).
8. Demetri, G. D. *et al.* Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. *N. Engl. J. Med.* **347,** 472–480 (2002).
9. Lynch, T. J. *et al.* Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. *N. Engl. J. Med.* **350,** 2129–2139 (2004).
10. Paez, J. G. *et al.* EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science* **304,** 1497–1500 (2004).
11. Sherr, C. J. Principles of tumor suppression. *Cell* **116,** 235–246 (2004).
12. Olivier, M. *et al.* The IARC TP53 database: new online mutation analysis and recommendations to users. *Hum. Mutat.* **19,** 607–614 (2002).
13. Bykov, V. J. *et al.* Restoration of the tumor suppressor function to mutant p53 by a low-molecular-weight compound. *Nature Med.* **8,** 282–288 (2002).
14. Snyder, E. L., Meade, B. R., Saenz, C. C. & Dowdy, S. F. Treatment of terminal peritoneal carcinomatosis by a transducible p53-activating peptide. *PLoS Biol.* **2,** E36 (2004).
15. Branda, C. S. & Dymecki, S. M. Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice. *Dev. Cell* **6,** 7–28 (2004).
16. Donehower, L. A. *et al.* Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356,** 215–221 (1992).
17. Jacks, T. *et al.* Tumor spectrum analysis in p53-mutant mice. *Curr. Biol.* **4,** 1–7 (1994).
18. Indra, A. K. *et al.* Temporally-controlled site-specific mutagenesis in the basal layer of the epidermis: comparison of the recombinase activity of the tamoxifen-inducible Cre-ER^{T} and Cre-ER^{T2} recombinases. *Nucleic Acids Res.* **27,** 4324–4327 (1999).
19. Soriano, P. Generalized *lacZ* expression with the ROSA26 Cre reporter strain. *Nature Genet.* **21,** 70–71 (1999).
20. Kemp, C. J., Wheldon, T. & Balmain, A. p53-deficient mice are extremely susceptible to radiation-induced tumorigenesis. *Nature Genet.* **8,** 66–69 (1994).
21. Collado, M. *et al.* Tumour biology: senescence in premalignant tumours. *Nature* **436,** 642 (2005).
22. Lowe, S. W. & Sherr, C. J. Tumor suppression by Ink4a-Arf: progress and puzzles. *Curr. Opin. Genet. Dev.* **13,** 77–83 (2003).
23. Xue, W. *et al.* Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* advance online publication, doi:10.1038/nature05529 (24 January 2007).
24. Zhang, L. *et al.* Wild-type p53 suppresses angiogenesis in human leiomyosarcoma and synovial sarcoma by transcriptional suppression of vascular endothelial growth factor expression. *Cancer Res.* **60,** 3655–3661 (2000).
25. Teodoro, J. G., Parker, A. E., Zhu, X. & Green, M. R. p53-mediated inhibition of angiogenesis through up-regulation of a collagen prolyl hydroxylase. *Science* **313,** 968–971 (2006).
26. Dickins, R. A. *et al.* Probing tumor phenotypes using stable and regulated synthetic microRNA precursors. *Nature Genet.* **37,** 1289–1295 (2005).
27. Christophorou, M. A. *et al.* Temporal dissection of p53 function *in vitro* and *in vivo*. *Nature Genet.* **37,** 718–726 (2005).
28. Christophorou, M. A., Ringshausen, I., Finch, A. J., Swigart, L. B. & Evan, G. I. The pathological response to DNA damage does not contribute to p53-mediated tumour suppression. *Nature* **443,** 214–217 (2006).
29. Vassilev, L. T. *et al. In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303,** 844–848 (2004).
30. Haupt, S. & Haupt, Y. Manipulation of the tumor suppressor p53 for potentiating cancer therapy. *Semin. Cancer Biol.* **14,** 244–252 (2004).
31. Olive, K. P. *et al.* Mutant p53 gain of function in two mouse models of Li-Fraumeni syndrome. *Cell* **119,** 847–860 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank N. Willis for helping to generate the LSL mice, H. Zheng for imaging mice, G. Wojtkiewicz for generating the movies with three-dimensional reconstruction, D. Crowley for help with histology, R. Bronson for reviewing the pathology, and M. Hemann for suggestions. A.V. is grateful to D. Ventura and G. Terranova for continuous support and encouragement. This work was supported by the Howard Hughes Medical Institute (T.J.), NCI (T.J., R.W., D.G.K.), and partially by a Cancer Center Support grant from the NCI (M.I.T.), the American Italian Cancer Research Foundation (A.V.), and the Leaf fund (D.G.K.). T.J. is the David H. Koch Professor of Biology and a Daniel K. Ludwig Scholar. D.A.T. is a Rita Allen Foundation Scholar.

Author Contributions A.V., D.G.K. and T.J. designed the experiments and wrote the paper. D.T. generated the p53-LSL mice and M.E.M. generated and characterized the Cre-ER^{T2} mice, determined the optimal Tamoxifen dosage, assisted with histopathological analysis and commented on the manuscript. A.V., D.G.K. and L.L. derived and characterized the tumour cell lines. A.V. performed the immunostainings, the TUNEL assays the SA-β-Gal stainings and the western blottings. A.V., D.K. and L.L. performed the tamoxifen intraperitoneal injections. E.E.R. derived the MEFs. L.L. and J.N. maintained the mouse colony and genotyped the animals. J.G. and D.G.K. evaluated the magnetic resonance images, J.G. supervised the magnetic resonance imaging, generated the three-dimensional reconstructions and determined tumour volumes. R.W. optimized *in vivo* imaging protocols, reviewed imaging data, discussed the results, and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.J. (tjacks@mit.edu).

LETTERS

Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript

Igor Martianov[1,2], Aroul Ramadass[1], Ana Serra Barros[1,3], Natalie Chow[1] & Alexandre Akoulitchev[1]

Alternative promoters within the same gene are a general phenomenon in gene expression[1,2]. Mechanisms of their selective regulation vary from one gene to another and are still poorly understood. Here we show that in quiescent cells the mechanism of transcriptional repression of the major promoter of the gene encoding dihydrofolate reductase depends on a non-coding transcript initiated from the upstream minor promoter and involves both the direct interaction of the RNA and promoter-specific interference. The specificity and efficiency of repression is ensured by the formation of a stable complex between non-coding RNA and the major promoter, direct interaction of the non-coding RNA with the general transcription factor IIB and dissociation of the preinitiation complex from the major promoter. By using *in vivo* and *in vitro* assays such as inducible and reconstituted transcription, RNA bandshifts, RNA interference, chromatin immunoprecipitation and RNA immunoprecipitation, we show that the regulatory transcript produced from the minor promoter has a critical function in an epigenetic mechanism of promoter-specific transcriptional repression.

The human gene encoding dihydrofolate reductase (*DHFR*) contains two promoters (Fig. 1a), with the downstream major promoter being responsible for the transcription of 99% of RNA[3]. Previous studies have suggested that the transcript generated from the upstream minor promoter could impede the formation of the preinitiation complex (PIC) on the major promoter *in vitro*[4]. In quiescent cells, the cell-cycle-regulated *DHFR* gene is repressed[5]. Indeed, the steady-state levels of *DHFR* mRNA are markedly reduced in a population of serum-starved, contact-inhibited U2OS cells (Fig. 1b, c). Under the same conditions quantitative real-time polymerase chain reaction (LCR) showed an accumulation of transcripts that initiated from the upstream minor *DHFR* promoter[4] and terminated within the second intron (Fig. 1d, e). After the example of the yeast *SER3* gene[2], we analysed the minor transcript for its role in the down regulation of the *DHFR* major promoter, possibly through RNA interference or RNA-mediated interactions.

We used a reconstituted transcription for the TATA-less *DHFR* gene[6] to compare transcription from the templates containing minor, major or both promoters in their original sequential order (Fig. 2a). By using the G-less cassette at the 3′ end of the transcriptional templates, we observed robust transcription from the major promoter, detectable transcription from the minor promoter and very low transcriptional activity when the minor promoter preceded the major one (Fig. 2a, lanes 1–3). This result suggested potential interference effect *in cis* between the two promoters under *in vitro* conditions.

Next we tested the effect of the non-coding transcript *in trans*. We reconstituted transcription from the major promoter in presence of either a separate transcriptional template generating the non-coding regulatory transcript, or the non-coding transcript itself. Transcription of the regulatory RNA *in trans* had a repressive effect on the major *DHFR* promoter (Fig. 2b, lanes 1–3). This effect was observed only when the generated transcript contained the sequence of the core major *DHFR* promoter. The template that stopped short of the core promoter did not have the same effect. This result indicates that the sequence corresponding to the core promoter is the essential part of the regulatory RNA. It also confirmed that the observed repression of the major promoter was not caused simply by promoter competition from the second template. We tested the repressive effect of the synthetic transcript directly. As shown in Fig. 2c (lanes 1–5), the addition of synthesized regulatory RNA (as generated from template B; Fig. 2b) but not the control RNA efficiently inhibited reconstituted transcription. These results strongly suggest that the mechanism of transcriptional repression may involve an RNA-mediated interaction.

Previous studies of transcription interference showed examples of the transcripts affecting various upstream activating sequences as part of the interference mechanism[2]. The Sp1-binding sites positioned between the minor and major promoters within the *DHFR* gene are indeed affected by the regulatory transcript. However, our *in vitro* study confirmed that transcription from the minimal core major promoter[6] could be suppressed *in trans* in the absence of Sp1 or any other upstream activating sequence. This indicated that one of the components of the reconstituted PIC might be a relevant target for the regulatory transcript. To check this hypothesis we analysed the recruitment of general transcription factors (GTFs) to the *DHFR* promoter. Two of the factors, TATA-binding protein (TBP) and TFIIB, have been shown to form a minimal stable core on many promoters in accordance with a stepwise assembly model[7]. Both factors show distinct properties for nucleic acid binding[8–10]. To monitor the effect of interfering transcription we applied quantitative chromatin immunoprecipitation (ChIP) to the reconstituted transcription (Fig. 3a). First, we tested the occupancy of TBP and TFIIB on a template containing only the major promoter. In the absence of initiated transcription we observed the recruitment of both factors to the major promoter. The addition of nucleoside triphosphates initiated transcription (Fig. 3a, lanes 2 and 4) without a significant change in the level of either TBP or TFIIB. In contrast, for the template containing both the minor and major promoters the levels of both factors decreased on the major promoter afer the induction of transcription with nucleoside triphosphates (Fig. 3b, lanes 2 and 4). The observed decrease was stoichiometric and consistent with the loss of the stable core TBP–TFIIB complex.

To test whether the transcript could interact directly with TBP or TFIIB, we used an RNA electrophoresis mobility-shift assay (EMSA). With ^{32}P-radiolabelled synthetic transcript as a probe, we observed no interaction with purified recombinant TBP or any bacterial proteins present in the extracts from the *Escherichia coli* strain used for

[1]Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK. [2]Institut de Genetique et de Biologie Moleculaire et Cellulaire, 1 rue Laurent Fries, BP 10142, 67404 Illkirch cedex, France. [3]Department of Biology, University of Aveiro, 3810-193 Aveiro, Portugal.

recombinant expression (Fig. 4a, lanes 2 and 3). Surprisingly, recombinant TFIIB formed a specific reversible ribonucleoprotein complex (Fig. 4a, b), suggesting a direct interaction.

To test whether the interaction with TFIIB could disrupt the PIC on the core major *DHFR* promoter we used a DNA bandshift assay (EMSA) developed earlier for the DHFR core promoter (Fig. 4c, lane 2)[6]. Titration of the synthetic minor transcript resulted in dissociation of the preformed PIC (Fig. 4c, lanes 1–5). In contrast, control RNA, as well as a non-specific competitor poly(dI-dC)·poly(dI-dC), had no affect (Fig. 4c, lane 6). Moreover, in agreement with the observed interaction between TFIIB and the regulatory transcript, the addition of recombinant TFIIB (rIIB) to the reconstituted transcription abolished inhibition by the synthetic regulatory RNA (Fig. 4d, lanes 1–4).

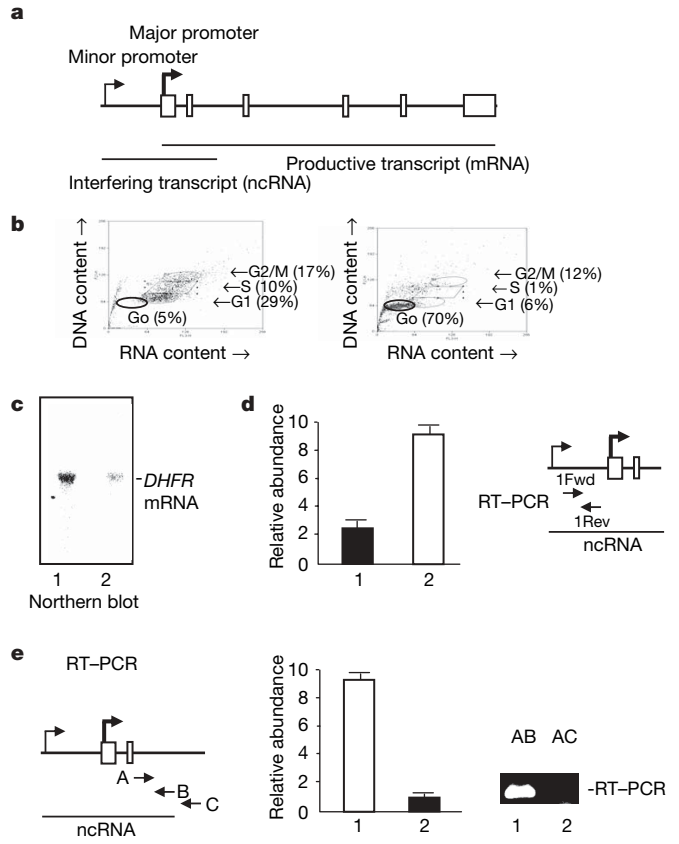


Figure 1 | Organization and regulation of the human *DHFR* promoter. **a**, Scheme of the promoter and exon organization of the *DHFR* gene. Previously described[3] productive (mRNA) and non-coding interfering (ncRNA) transcripts are indicated. Boxes represent exons. **b**, FACS sorting of the U2OS cells used in the study. FACS diagrams of cells grown in the presence of 10% fetal calf serum (left, cycling cells) and under contact inhibition in presence of 0.5% fetal calf serum (right, quiescent cells) are shown. The percentages of G0, G2/M, S and G1 cells under each set of conditions are indicated. **c**, Northern blot analysis of *DHFR* mRNA. Lane 1 shows total RNA from cycling cells (see **b**), and lane 2 shows total RNA from quiescent cells (see **b**). DNA corresponding to four, five and six exons was used as a probe. **d**, Real-time RT–PCR analysis of minor ncRNA transcript in cycling (1, see also **b**) and quiescent (2, see also **b**) cells. The values shown are calculated from three independent experiments. Transcript was detected between the minor and major transcription start sites. The major and minor promoters are symbolized as in **a**. **e**, Termination of *DHFR* minor transcripts in quiescent cells. RT–PCR was designed to evaluate the levels of transcription in quiescent cells before (lane 1) and after (lane 2) the reported termination site within the second intron of the DHFR gene. Products of RT–PCR were separated on an agarose gel (lane 1, product of RT–PCR with primers A and B; lane 2, product of RT–PCR with primers A and C). The major and minor promoters are symbolized as in **a**. Error bars indicate s.d.

We then performed an RNA immunoprecipitation assay[11] and showed that immunoprecipitation of TFIIB from quiescent cells contained regulatory transcript (Fig. 4e, lane 2 and 9). This strongly suggested that the interaction of TFIIB with the transcript also occurs *in vivo*.

To confirm that disruption of the PIC could also take place *in vivo* we conducted quantitative ChIP of TBP, TFIIB and RNA polymerase II in U2OS cells under conditions of normal growth and quiescence. We observed a marked decrease in the occupancy of the major *DHFR* promoter by all three factors in quiescent cells (Fig. 4f, lanes 1–6). In contrast, the occupancy of the minor promoter remained unchanged (Fig. 4f, lanes 7–12).

Next, we tested whether degradation of the regulatory transcript by RNA interference would preserve the PIC on the major promoter (Fig. 4h). Cells were treated with control small interfering RNA (siRNA) or with siRNA targeting the regulatory transcript (Fig. 4h, lanes 1 and 2). Reverse-transcriptase-mediated PCR (RT–PCR) and ChIP assays indicated that when the regulatory transcript was specifically degraded, the occupancy of TFIIB on the major promoter remained high. These results clearly implicated the non-coding regulatory transcript in disruption of the PIC on the major *DHFR* promoter *in vivo*.

To strengthen the evidence for RNA-dependent repression of the major *DHFR* promoter *in vivo* we established two inducible systems. First, we replaced the minor promoter with the inducible human metallothionein promoter[12] and analysed the occupancy of TFIIB on the major promoter by ChIP. We observed a decrease in occupancy of TFIIB on the *DHFR* major promoter within 2 h after induction (Fig. 3c). The conditions of the induction itself had no effect on either the *DHFR* promoter or any other unrelated control

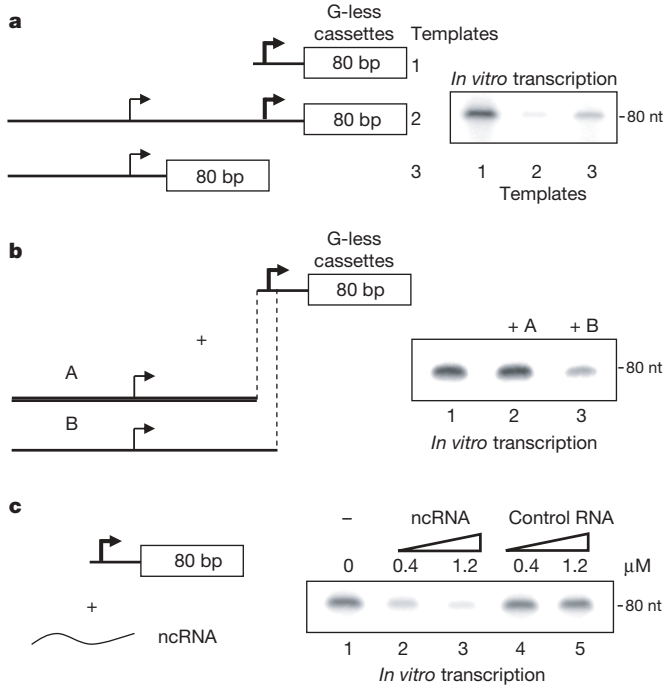


Figure 2 | Transcriptional interference *in vitro*. **a**, Inhibition of *in vitro* transcription from the major promoter *in cis*. Transcription from templates 1, 2 and 3 was reconstituted *in vitro* (lanes 1, 2 and 3, respectively). Each template contained an 80-base-pair (bp) G-less cassette as shown. **b**, Inhibition of *in vitro* transcription from the major promoter *in trans*. Transcription was reconstituted on template 1 (**a**) (lane 1) in the presence of the template containing the minor promoter, and extended either to (template A) or past (template B) the core *DHFR* promoter sequence. **c**, Inhibition of *in vitro* transcription from the major promoter *in trans*. Transcription was reconstituted on template 1 (**a**) (lane 1) in the presence of the synthetic non-coding DHFR transcript (lanes 2 and 3; **b**, template B) or control RNA (lanes 4 and 5). nt, nucleotides.

promoters[12] (data not shown). Second, we deployed the Tet-dependent inducible T-REx system[13], in which the minor promoter was replaced by the cytomegalovirus (*CMV*) promoter under the control of the Tet repressor. We analysed steady-state pools of Tet-repressed and CMV-active U2OS cells for the occupancy of TFIIB on the major *DHFR* promoter. Under the active *CMV* promoter (ON condition in Fig. 3d)—that is, when the non-coding transcript was synthesized—we observed a loss of TFIIB on the major *DHFR* promoter (Fig. 3d). To exclude the direct interference of regulatory protein complexes we increased the distance between the promoters by inserting more than 2 kilobases of sequence that did not contain any pausing or terminating sites (Fig. 3d). On induction of transcription from the upstream promoter we observed a similar loss of TFIIB on the major promoter to that in the original construct.

We further analysed the role of transcriptional interference by inserting a termination sequence between two promoters (Fig. 3d). On induction of the upstream promoter (ON state), the occupancy of TFIIB on the major promoter remained high in the presence of the terminator. These data strongly support a direct role for the regulatory transcript in the repression of the major *DHFR* promoter.

To address the specificity of the RNA-dependent repression we investigated the properties of the major *DHFR* promoter, which is GC-rich and contains several G-track sequences. Such sequences have been shown to form stable purine-purine-pyrimidine triplex structures (H form) between DNA and RNA[14–16]. It was shown previously that the triplex-forming oligonucleotides can inhibit the transcription of the *Ets2* gene in prostate cancer cells[17]. The Sp1-binding site upstream of the major *DHFR* promoter start site has been reported to form a stable ($K_d \approx 1\ \mu M$) triplex complex[18]. We analysed the formation of the stable triplex structure between the major *DHFR* promoter and the regulatory transcript. A standard H-form band-shift assay revealed the formation of a stable specific DNA–RNA complex (Fig. 4g, lanes 2–5). To increase the resolution of the assay we tested a synthetic oligoribonucleotide corresponding to the core

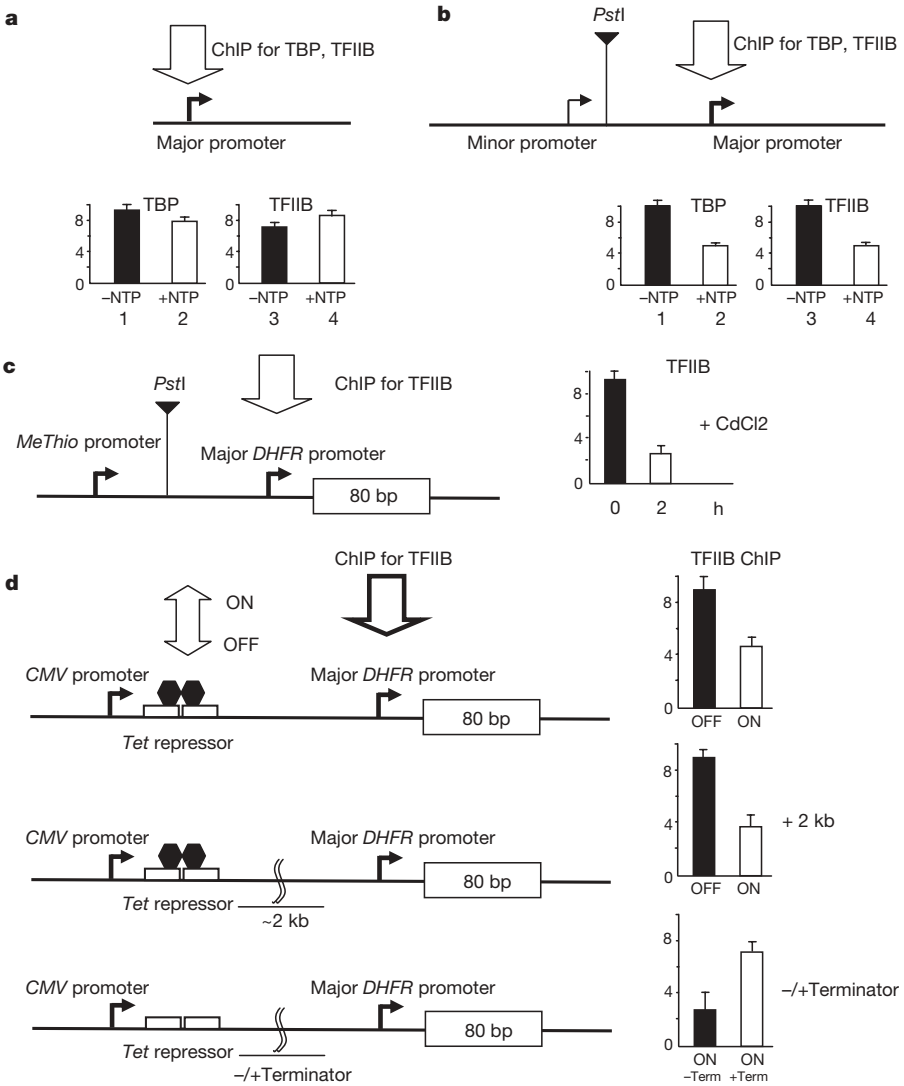


Figure 3 | Dissociation of TBP and TFIIB from the major *DHFR* promoter under induced transcriptional interference. **a**, **b**, *In vitro* occupancy for TBP and TFIIB on the major *DHFR* promoter in a reconstituted system after formation of the preinitiation complex (−NTP) (**a**, lanes 1 and 3; **b**, lanes 1 and 3) and after the initiation of transcription (+NTP) (**a**, lanes 2 and 4; **b**, lanes 2 and 4). Quantitative immunoprecipitations were performed on templates containing only the major promoter (**a**, lanes 1–4) and both the minor and major promoters (**b**, lanes 1–4). The *Pst*I restriction site was used for efficient resolution between the minor and major promoters in quantitative real-time PCR after ChIP immunoprecipitation (see Methods). **c**, Relative occupancy by TFIIB on the major *DHFR* promoter placed downstream of the human *MeThio* promoter. Occupancy is shown before and 2 h after induction of the *MeThio* promoter. **d**, Occupancy by TFIIB of the *DHFR* major promoter placed downstream of the tetracycline-regulated *CMV* promoter (see Methods). Occupancy is shown under conditions of repressed (OFF) and active (ON) *CMV* promoter on the construct with the original distance between the promoters, with a 2-kilobase (kb) insertion between the promoters and with the termination sequence inserted between the promoters (see Methods). Error bars indicate s.d.

of the major promoter only. A stable H form between labelled promoter DNA and the oligoribonucleotide was also detected, with the K_d being similar to values reported previously[18] (Fig. 4g, lanes 9–12).

Our analysis of an alternative promoter revealed an RNA-dependent mechanism of transcriptional repression. It is a dynamic process dependent on the production of the regulatory transcript from the upstream start site. Highly specific and stable triplex structures formed by the transcript with the major promoter may well contribute to promoter targeting and repression *in vivo*[17] (Fig. 4h). The transcript can interact with basal TFIIB and dissociate the PIC efficiently. This interaction is particularly intriguing in view of the inducible regulatory conformational switch in TFIIB[19].

These regulatory interactions should be viewed in the full context of the epigenetic mechanism of repression as it takes place on the *DHFR* promoter *in vivo*. Preliminary analysis (A.A., unpublished observations) has revealed similarities to epigenetic markers associated with transcriptional gene silencing induced by the targeted application of synthetic antigene RNAs complementary to the

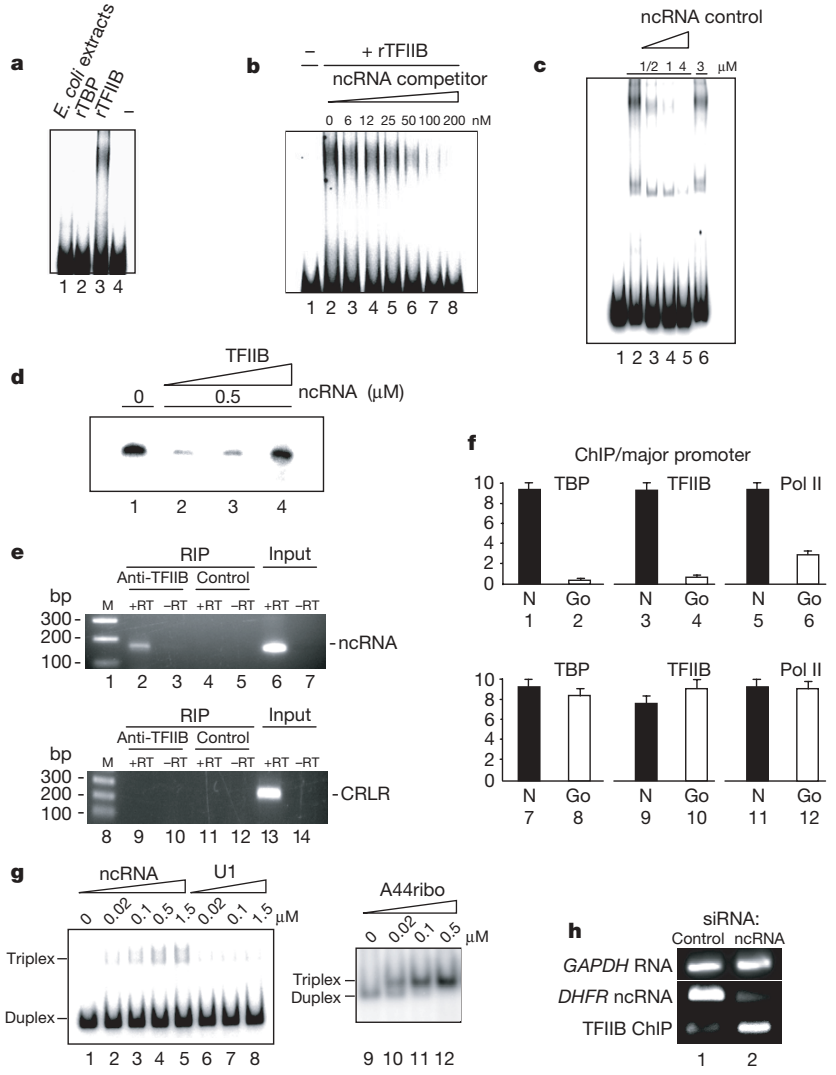


Figure 4 | Interactions with the regulatory transcript *in vitro* and *in vivo*. **a**, RNA EMSA of recombinant TBP (rTBP; lane 3) and recombinant TFIIB (rTFIIB; lane 4) with the non-coding minor transcript (lane 1). Lane 2 contained extracts from the *E. coli* BL21 strain used for the expression of rTBP and rTFIIB. **b**, RNA EMSA of the specific rTFIIB complex (lanes 2–8) in the presence of the specific competitor, non-coding *DHFR* minor transcript (lanes 3–8). Lane 1 shows free probe. **c**, DNA EMSA of the PIC on the minimal core of the major *DHFR* promoter. The bandshift was performed as described previously[6]. After formation of the PIC, non-coding *DHFR* minor transcript was added as a specific competitor (lanes 1–5). Lane 6 shows competition with control RNA. **d**, Rescue of inhibition by rTFIIB *in vitro*. *In vitro* transcription was run in the absence (lane 1) or in the presence (lanes 2–4) of regulatory RNA. Increasing concentrations of rTFIIB were added to reactions 2, 3 and 4. **e**, Association of non-coding *DHFR* minor transcript with TFIIB *in vivo*. RNA immunoprecipitation was performed with anti-TFIIB antibody. Co-precipitated *DHFR* minor transcript (lane 2) was detected with the specific primers shown in Fig. 1d. No control calcitonin receptor-like receptor (*CRLR*) mRNA was co-precipitated with anti-TFIIB antibody (lane 9). Lanes 6 and 13 represent the input, −RT represents the control for DNA contamination. M, DNA size marker; RT, reverse transcription. **f**, *In vivo* occupancy for TBP, TFIIB and DNA polymerase II (Pol II) on the major *DHFR* promoter (lanes 1–6) or in the major *DHFR* promoter (lanes 7–12). Quantitative immunoprecipitation was performed with cycling (lanes 1, 3, 5, 7, 9 and 11) and quiescent (lanes 2, 4, 6, 8, 10 and 12) cells. The values shown represent three independent experiments. Error bars indicate s.d. **g**, Triplex complex formation between *DHFR* major promoter and regulatory RNA. Formation of the triple helix was monitored with a ^{32}P-labelled fragment of 140-bp DNA and regulatory RNA of the same size (lanes 2–5) or U1 snRNA as a non-specific control (lanes 6–8). Formation of the H form was also assayed between the ^{32}P-labelled 30-bp DNA duplex (lane 9) and increasing amounts of the oligoribonucleotide (A44ribo) corresponding to the core of the major promoter (lanes 10–12; see Methods). **h**, ncRNA maintains low occupancy of TFIIB on the major *DHFR* promoter *in vivo*. RT–PCR analysis of DHFR ncRNA and control glyceraldehyde-3-phosphate dehydrogenase (GAPDH) RNA in cells incubated with siRNA against *DHFR* ncRNA (lane 2) or control siRNA (lane 1). Cells with intact (lane 1) or degraded (lane 2) ncRNA were analysed by ChIP assay for TFIIB occupancy on the major *DHFR* promoter.

promoter regions of mammalian genes[20,21]. Further analysis of TFIIB-mediated repression should provide details of the epigenetic mechanism of promoter regulation by a non-coding RNA.

METHODS

Details of fluorescence-activated cell sorting (FACS), reconstituted transcription *in vitro*, EMSAs, RNA interference, ChIP and quantitative immunoprecipitation from the *in vitro* system, RNA immunoprecipitation, H-form formation, and inducible transcription are given in Supplementary Information.

Received 14 September; accepted 8 December 2006.
Published online 21 January 2007.

1. Ayoubi, T. A. & Van De Ven, W. J. Regulation of gene expression by alternative promoters. *FASEB J.* **10,** 453–460 (1996).
2. Martens, J. A., Laprade, L. & Winston, F. Intergenic transcription is required to repress the *Saccharomyces cerevisiae SER3* gene. *Nature* **429,** 571–574 (2004).
3. Masters, J. N. & Attardi, G. Discrete human dihydrofolate reductase gene transcripts present in polysomal RNA map with their 5′ ends several hundred nucleotides upstream of the main mRNA start site. *Mol. Cell. Biol.* **5,** 493–500 (1985).
4. Blume, S. W., Meng, Z., Shrestha, K., Snyder, R. C. & Emanuel, P. D. The 5′-untranslated RNA of the human dhfr minor transcript alters transcription pre-initiation complex assembly at the major (core) promoter. *J. Cell. Biochem.* **88,** 165–180 (2003).
5. Hendrickson, S. L., Wu, J. S. & Johnson, L. F. Cell cycle regulation of dihydrofolate reductase mRNA metabolism in mouse fibroblasts. *Proc. Natl Acad. Sci. USA* **77,** 5140–5144 (1980).
6. Akoulitchev, S., Makela, T. P., Weinberg, R. A. & Reinberg, D. Requirement for TFIIH kinase activity in transcription by RNA polymerase II. *Nature* **377,** 557–560 (1995).
7. Orphanides, G., Lagrange, T. & Reinberg, D. The general transcription factors of RNA polymerase II. *Genes Dev.* **10,** 2657–2683 (1996).
8. Kim, J. L., Nikolov, D. B. & Burley, S. K. Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature* **365,** 520–527 (1993).
9. Nikolov, D. B. *et al.* Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377,** 119–128 (1995).
10. Lagrange, T. *et al.* High-resolution mapping of nucleoprotein complexes by site-specific protein-DNA photocrosslinking: organization of the human TBP-TFIIA-TFIIB-DNA quaternary complex. *Proc. Natl Acad. Sci. USA* **93,** 10620–10625 (1996).
11. Gilbert, C., Kristjuhan, A., Winkler, G. S. & Svejstrup, J. Q. Elongator interactions with nascent mRNA revealed by RNA immunoprecipitation. *Mol. Cell* **14,** 457–464 (2004).
12. Kazmi, M. A., Dubin, R. A., Oddoux, C. & Ostrer, H. High-level inducible expression of visual pigments in transfected cells. *Biotechniques* **21,** 304–311 (1996).
13. Xu, Z. L., Mizuguchi, H., Mayumi, T. & Hayakawa, T. Regulated gene expression from adenovirus vectors: a systematic comparison of various inducible systems. *Gene* **309,** 145–151 (2003).
14. Morgan, A. R. & Wells, R. D. Specificity of the three-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. *J. Mol. Biol.* **37,** 63–80 (1968).
15. Letai, A. G., Palladino, M. A., Fromm, E., Rizzo, V. & Fresco, J. R. Specificity in formation of triple-stranded nucleic acid helical complexes: studies with agarose-linked polyribonucleotide affinity columns. *Biochemistry* **27,** 9108–9112 (1988).
16. Marck, C. & Thiele, D. Poly(dG).poly(dC) at neutral and alkaline pH: the formation of triple stranded poly(dG).poly(dG).poly(dC). *Nucleic Acids Res.* **5,** 1017–1028 (1978).
17. Carbone, G. M., McGuffie, E. M., Collier, A. & Catapano, C. V. Selective inhibition of transcription of the Ets2 gene in prostate cancer cells by a triplex-forming oligonucleotide. *Nucleic Acids Res.* **31,** 833–843 (2003).
18. Gee, J. E., Blume, S., Snyder, R. C., Ray, R. & Miller, D. M. Triplex formation prevents Sp1 binding to the dihydrofolate reductase promoter. *J. Biol. Chem.* **267,** 11163–11167 (1992).
19. Roberts, S. G. & Green, M. R. Activator-induced conformational change in general transcription factor TFIIB. *Nature* **371,** 717–720 (1994).
20. Janowski, B. A. *et al.* Involvement of AGO1 and AGO2 in mammalian transcriptional silencing. *Nat. Struct. Mol. Biol.* **13,** 787–791 (2006).
21. Kim, D. H., Villeneuve, L. M., Morris, K. V. & Rossi, J. J. Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. *Nat. Struct. Mol. Biol.* **13,** 793–797 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank members of the laboratory for advice and support, N. Rust for providing support in flow cytometry analysis, S. Roberts and T. Oelgeschlaeger for suggestions, and J. Svejstrup for advice on the RNA immunoprecipitation procedure. This work was supported by grants from the Wellcome Trust and Medical Research Council. I.M. was supported by EMBO long-term fellowship.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.A. (alexandre.akoulitchev@path.ox.ac.uk).

LETTERS

Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation

Guillermo Barreto[1]*, Andrea Schäfer[1]*, Joachim Marhold[2], Dirk Stach[2], Suresh K. Swaminathan[1], Vikas Handa[1], Gabi Döderlein[1], Nicole Maltry[1], Wei Wu[1]†, Frank Lyko[2] & Christof Niehrs[1]

DNA methylation is an epigenetic modification that is essential for gene silencing and genome stability in many organisms. Although methyltransferases that promote DNA methylation are well characterized, the molecular mechanism underlying active DNA demethylation is poorly understood and controversial[1,2]. Here we show that Gadd45a (growth arrest and DNA-damage-inducible protein 45 alpha), a nuclear protein involved in maintenance of genomic stability, DNA repair and suppression of cell growth[3,4], has a key role in active DNA demethylation. *Gadd45a* overexpression activates methylation-silenced reporter plasmids and promotes global DNA demethylation. *Gadd45a* knockdown silences gene expression and leads to DNA hypermethylation. During active demethylation of *oct4* in *Xenopus laevis* oocytes[5], Gadd45a is specifically recruited to the site of demethylation. Active demethylation occurs by DNA repair and Gadd45a interacts with and requires the DNA repair endonuclease XPG. We conclude that Gadd45a relieves epigenetic gene silencing by promoting DNA repair, which erases methylation marks.

To identify proteins capable of promoting active DNA demethylation, we screened a *Xenopus* expression library for complementary DNAs able to re-activate the transcription of a methylation-silenced luciferase reporter gene. Sib selection of three independent active pools each led to the isolation of *Gadd45a*. *Gadd45* genes are stress inducible and function in numerous biological processes, for example, the cell cycle, senescence, apoptosis and nucleotide excision repair[4].

Transfection of *Gadd45a* activated all methylated luciferase (Fig. 1a, d; Supplementary Fig. 1) and green fluorescent protein (GFP) (Fig. 1c) reporter plasmids tested, and did so in various cell lines, indicating that its effect is promoter-, reporter- and cell-type-independent. *Gadd45a* also activated unmethylated reporter plasmids (Fig. 1b; Supplementary Fig. 1). However, unmethylated plasmid DNA is known to become *de novo* methylated following transient transfection[6]. Indeed, treatment of HEK293T cells with the DNA methyltransferase inhibitor 5-aza-2'-deoxycytidine (AZA) activated unmethylated reporters similarly to *Gadd45a* (Fig. 1b), whereas it did not affect methylated reporters (Supplementary Fig. 1).

Because a hallmark of Gadd45 overexpression is its ability to induce growth arrest[4], we tested if blocking cell proliferation activates the reporter. Neither transfection with the cell cycle inhibitor p28 (ref. 7; Fig. 1a, b; Supplementary Fig. 1), nor aphidicolin treatment

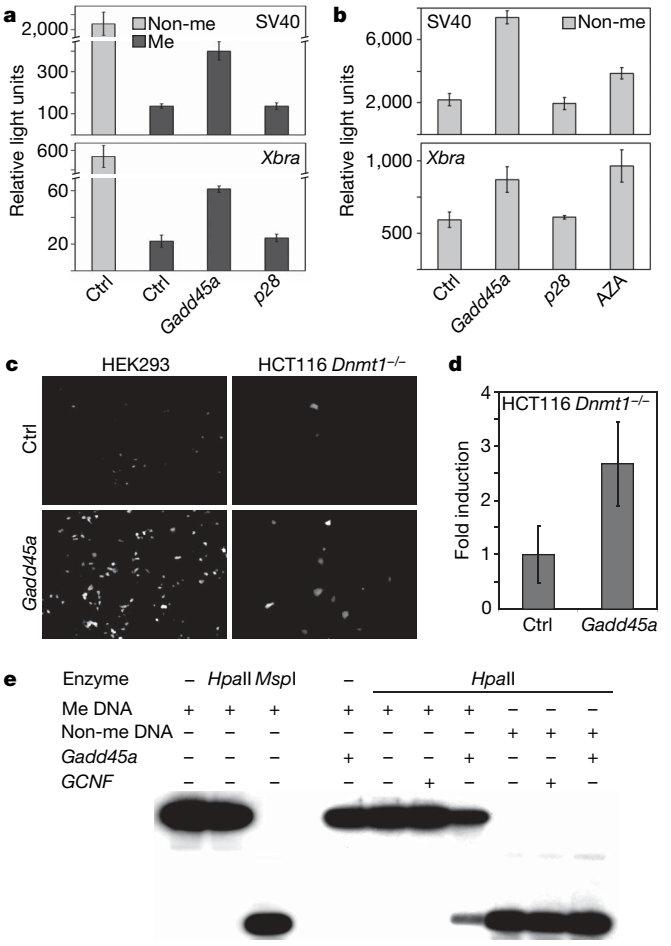


Figure 1 | *Gadd45a* activates methylation-silenced reporter plasmids. **a**, Luciferase reporter assays of HEK293T cells transiently transfected with pBS (Ctrl) or the indicated genes. Reporters containing the SV40- or the *Xenopus brachyury* promoter (*Xbra*) were used. Reporter plasmids were unmethylated (Non-me) or *in vitro* methylated (Me) with M.SssI methylase. Error bars, s.e.m. ($n = 3$). **b**, Luciferase reporter assays of HEK293T cells as in **a**. AZA, 5-aza-2'-deoxycytidine treatment. Error bars, s.e.m. ($n = 3$). **c**, *Gadd45a* activation of methylated reporters is *Dnmt1*-independent. Fluorescence microscopy of HEK293 and HCT116 *Dnmt1*$^{-/-}$ cells transiently transfected with methylated *pCS2–GFP* reporter and pBS (Ctrl) or *Gadd45a*. **d**, Luciferase reporter assay of cells transiently transfected with methylated SV40 reporter. Error bars, s.e.m. ($n = 3$). **e**, *Gadd45a* overexpression induces demethylation of a mouse *oct4* promoter plasmid. Methylated reporter plasmid *pOctTK–GFP* was recovered from HEK293 cells after transient transfection with *Gadd45a* or germ cell nuclear factor (*GCNF*, control). Recovered plasmids were digested with the indicated enzyme and the products analysed by Southern blot using a *GFP* probe.

[1]Division of Molecular Embryology, [2]Division of Epigenetics, German Cancer Research Center, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany. †Present address: Department of Biological Sciences and Biotechnology, Tsinghua University, 100084 Beijing, China.
*These authors contributed equally to this work.

(not shown) activated the reporter, arguing against a proliferation-dependent effect. Furthermore, although *Gadd45a* expression induced some G1 arrest and apoptosis in HEK293 cells, it had no marked effects on HeLa or HEK293T cells, or *Xenopus* oocytes (not shown).

The methylation state of methylated reporter plasmid containing the 5′ upstream region of the pluripotency marker *oct4* (ref. 8) was analysed by methylation-sensitive restriction enzyme cleavage and Southern blotting (Fig. 1e). *In vitro* methylated *oct4* reporter plasmid was digested by *Msp*I, but was resistant to the methylation-sensitive isoschizomer *Hpa*II, when recovered from control (GCNF)-transfected cells, indicating that the methylation status was not altered. In contrast, plasmid DNA recovered from *Gadd45a*-transfected cells was digested by *Hpa*II, indicating that it had undergone demethylation.

We next turned to endogenous *oct4*, which is silenced by DNA methylation in somatic cells, but can be activated by the methylation inhibitor AZA in combination with the histone deacetylase inhibitor trichostatin A (TSA)[9]. Our results confirm that *oct4* expression can be induced by combined AZA/TSA treatment as shown by PCR after reverse transcription (RT–PCR) in NIH/3T3 cells (Fig. 2a). Furthermore, AZA could be substituted by *Gadd45a* transfection, suggesting that *Gadd45a* promotes demethylation of endogenous *oct4*. This suggests that DNA demethylation and histone acetylation act in concert to activate the *oct4* promoter.

To confirm demethylation, we carried out bisulphite sequencing of *oct4*. Demethylation of its upstream region is associated with nuclear reprogramming in *Xenopus*[5]. The methylation pattern of the ten CpGs analysed in NIH/3T3 cells showed that the proximal-most six CpGs are predominantly methylated, whereas the distal CpGs show less methylation, as has been shown in other cell types[5,10]. *Gadd45a* transfection induced complete *oct4* demethylation in 20% of the clones analysed (Fig. 2b) and reduced the methylation level in several other clones. The persistence of high methylation levels in 60–70% of the clones is most likely due to the low transfection efficiency of 3T3 cells (30%, as tested with GFP). Proximal CpGs were predominantly demethylated (numbers 4–10, Fig. 2b), which was confirmed in *Xenopus* (see below). The results indicate that *Gadd45a* can induce demethylation of single copy genes. Furthermore, *Gadd45a* overexpression induced hypomethylation of chromosome 1 satellite 2 (C1S2). Combined bisulphite restriction analysis (COBRA) showed a significant increase in unmethylated C1S2 (Fig. 2c), indicating that Gadd45 activity is not restricted to single-copy genes but acts as well on repetitive elements.

To characterize further the effect of *Gadd45a* on global DNA methylation we quantified total 5-methylcytosine (5mC) levels by capillary electrophoresis (Fig. 2d). In HEK293 cells, AZA treatment led to a threefold 5mC reduction, from 2.1 to 0.7% in dividing cells, but not in serum-starved, non-proliferating cells, and so is indicative of passive demethylation. In contrast, *Gadd45a* transient transfection led to a similar reduction of 5mC to 0.9% in both dividing as well as non-proliferating cells, demonstrating active demethylation. We also tested 3T3 cells and because of their low transfection efficiency, we used immunostaining with 5mC-specific antibodies. Only cells positive for GFP–Gadd45a fusion protein expression were scored. *GFP–Gadd45a* transfection reduced the number of strongly 5mC positive cells from 91% ($n = 842$) to 43% ($n = 819$) (Fig. 2e), which confirms that *Gadd45a* expression induces global DNA hypomethylation.

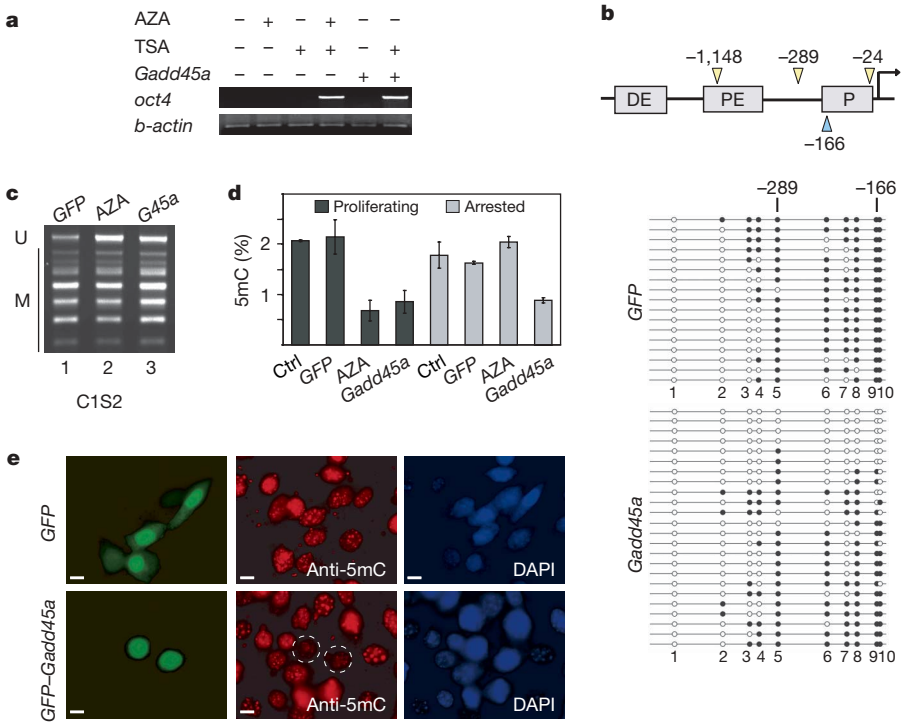


Figure 2 | *Gadd45a* promotes gene-specific and global DNA demethylation. **a**, NIH/3T3 cells were treated with AZA and/or trichostatin A (TSA) and/or transiently transfected with *Gadd45a*. Expression of *oct4* was monitored by RT–PCR. **b**, Diagram of mouse *oct4* (top panel). P, promoter; PE, proximal enhancer; DE, distal enhancer. Four restriction sites for methyl-sensitive enzymes (arrowheads; yellow, *Hpa*II; blue, *Hha*I) used for methyl-sensitive PCR (Fig. 3e, f) are indicated. The position of four CpGs is indicated. Bisulphite sequence analysis of *oct4* promoter in NIH/3T3 cells transiently transfected with *GFP* or *Gadd45a* (lower panel). White and black circles, unmethylated and methylated CpG, respectively. **c**, *Gadd45a* overexpression induces demethylation of repetitive sequences. Methylation of chromosome 1 satellite 2 (C1S2) in AZA-treated or transiently transfected HCT116 cells was analysed by COBRA. U, undigested (unmethylated) PCR amplicon; M, digested (methylated) restriction fragments. **d**, Capillary electrophoresis of DNA 5-methylcytosine (5mC). HEK293 cells were either treated with AZA or transiently transfected. Cells were either proliferating or arrested by serum starvation. After 48 h 5mC levels were determined by capillary electrophoresis. Error bars, s.e.m. ($n = 5$). **e**, Immunostaining of transiently transfected NIH/3T3 cells with 5mC specific antibody (anti-5mC). Note the weak anti-5mC staining in two encircled GFP–Gadd45a-positive cells. Scale bar, 10 μm.

Specific short interfering RNA (siRNA) knockdown of *Gadd45a* and *Gadd45b* was accompanied by downregulation of *MLH1*, *THBS1* and *p16* (Fig. 3a), three genes known to be regulated by DNA methylation[11–13]. To confirm that downregulation was due to increased DNA methylation, we analysed the *MLH1* promoter using COBRA (Fig. 3b). In untreated cells, *MLH1* is fully methylated between −921 to −446, whereas the two proximal regions are partially methylated. Knockdown of *Gadd45a* and *Gadd45b* induces complete methylation of these proximal regions. We conclude that *Gadd45* is required to keep *MLH1* partially demethylated. In agreement with the COBRA data, knockdown of *Gadd45a* and *Gadd45b* either individually or in combination, increased total 5mC levels twofold (Fig. 3c) and induced hypermethylation of LINE elements and pericentric satellite repeat sequences (Supplementary Fig. 2). We therefore conclude that *Gadd45a* is required for global DNA-demethylation as well as single copy gene demethylation.

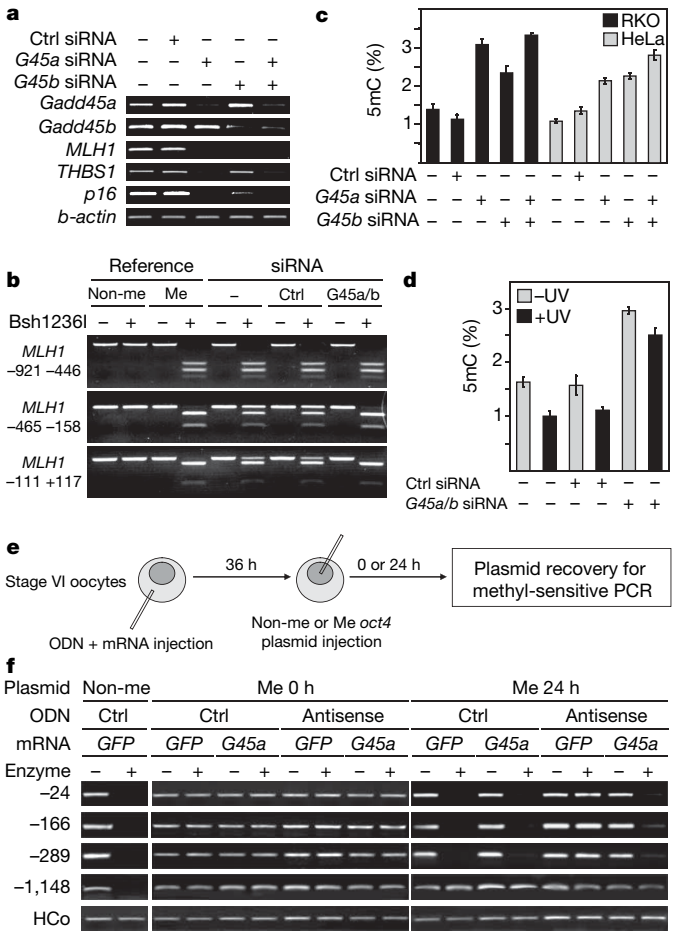


Figure 3 | *Gadd45a* loss of function induces DNA hypermethylation. **a**, RKO cells were transfected with *Gadd45* isoform-specific siRNAs and RT–PCR analysis performed for the indicated genes after 48 h. G45, Gadd45. **b**, *Gadd45* knockdown causes DNA hypermethylation of *MLH1*. COBRA of RKO cells transfected with indicated siRNAs as in **a**. Reference, COBRA of *in vitro* methylated and unmethylated control DNA. *Bsh*12361, restriction enzyme used for diagnostic digest. **c**, *Gadd45* knockdown induces global DNA hypermethylation. RKO and HeLa cells were transfected with the indicated siRNAs as in **a**. Capillary electrophoresis analysis of DNA 5mC was performed as in Fig. 2d. Error bars, s.e.m. ($n = 4$). **d**, UV-induced DNA hypomethylation is *Gadd45*-dependent. RKO cells were transfected with the indicated siRNAs as in **a** and UV irradiated. Capillary electrophoresis analysis of DNA 5mC was performed as in Fig. 2d. Error bars, s.e.m. ($n = 4$). **e**, Active DNA demethylation of the *oct4* promoter in *Xenopus* oocytes requires *Gadd45a*. Diagram of the experimental procedure. Control (Ctrl) or antisense oligodeoxyribonucleotides (ODN) and *GFP*- or *Gadd45a* mRNA were injected into the cytoplasm of stage VI oocytes. After 36 h, oocytes were injected into the nucleus with unmethylated (Non-me) or *Hpa*II/*Hha*I methylated (Me) *oct4* plasmid. **f**, PCR analysis of four CpGs indicated in Fig. 2b with (+) and without (−) methyl-sensitive restriction enzyme digestion of *oct4* plasmids as described in **d**. HCo, *oct4* amplicon lacking *Hpa*II/*Hha*I sites as loading control.

Because DNA methylation has been previously linked to nuclear reprogramming[14], we tested if *Gadd45a* is also involved in active DNA demethylation that occurs in *Xenopus* oocytes. *Xenopus* oocytes actively demethylate an *in vitro* methylated mouse *oct4* construct within 24 h (ref 5). Using this model system, stage VI oocytes were injected with antisense oligodeoxynucleotides (ODN)[15] and incubated for 36 h, to allow depletion of maternal *Gadd45a* messenger RNA and translation of the co-injected mRNAs to proceed (Supplementary Fig. 3a–e). Oocytes were re-injected with methylated *oct4* plasmid into the nucleus and either harvested immediately (0 h) or, to allow demethylation to proceed, after 24 h (Fig. 3e). Finally, *oct4* plasmid was analysed for the methylation status of four CpGs located at −24, −166, −289 and −1,148 (Fig. 2b) using methyl-sensitive enzyme restriction coupled to PCR amplification. When *in vitro* methylated *oct4* plasmid was recovered at 0 h it was resistant to methylation-specific restriction enzyme digestion at all four CpG sites and under all conditions (Fig. 3f). However, 24 h after injection, site-specific *oct4* demethylation had occurred at −24, −166 and −289, but not at the distal site, −1,148, as reported previously[5]. Importantly, in oocytes depleted of *Gadd45a*, this demethylation was abolished. Inhibition was reversed by injection of excess human *Gadd45a* mRNA, confirming the specificity of the depletion. Like in 3T3 cells (Fig. 2b), human *Gadd45a* mRNA injection demethylated the proximal CpGs, but not site −1,148. These results demonstrate that *Gadd45a* is required for site-specific demethylation of *oct4* in *Xenopus* oocytes.

Demethylation of 5-methylcytosine (5mC) residues may occur by an active mechanism or a passive mechanism involving DNA replication. However, the Gadd45-mediated demethylation is active, because it occurred in non-dividing oocytes as well as in non-proliferating cells (Fig. 2d). Additional evidence against a passive demethylation process are the observations that: (1) *Gadd45a* activated methylated reporters whereas the methyltransferase inhibitor AZA does not (Supplementary Fig. 1); (2) *Gadd45a* activated methylated reporters in HCT116 *Dnmt1*$^{-/-}$ cells[16], which are devoid of the maintenance methyltransferase (Fig. 1c, d); and (3) *Gadd45a* induced DNA demethylation of a non-replicating plasmid (Fig. 1e).

Gadd45a has been implicated in nucleotide excision repair[17,18]. Because DNA repair has previously been linked to active DNA demethylation[19–22], this suggested an attractive mechanism for *Gadd45a*-mediated demethylation. To test for DNA synthesis, characteristic of DNA repair, bromodeoxyuridine (BrdU) was co-injected into oocytes with methylated or unmethylated *oct4* plasmid (Supplementary Fig. 4). After 12 or 36 h, plasmid DNA was immunoprecipitated with anti-BrdU antibodies and analysed by PCR. A PCR product increase was observed with time (Fig. 4a), reflecting progressive BrdU incorporation, and this was inhibited by *Gadd45a* depletion. No BrdU-labelled *oct4* was detected if the injected plasmid was unmethylated. We conclude that DNA synthesis accompanies active DNA demethylation of *oct4* in *Xenopus* oocytes and that this synthesis requires *Gadd45a*.

To corroborate a direct involvement of *Gadd45a*, we carried out chromatin immunoprecipitation (ChIP). In *Xenopus* oocytes, Gadd45a was specifically targeted to the *oct4* promoter (Fig. 4b), encompassing the demethylated sites −166 and −289, but not to the proximal enhancer, or to *GFP* coding region. Promoter binding occurred on both methylated as well as unmethylated *oct4*.

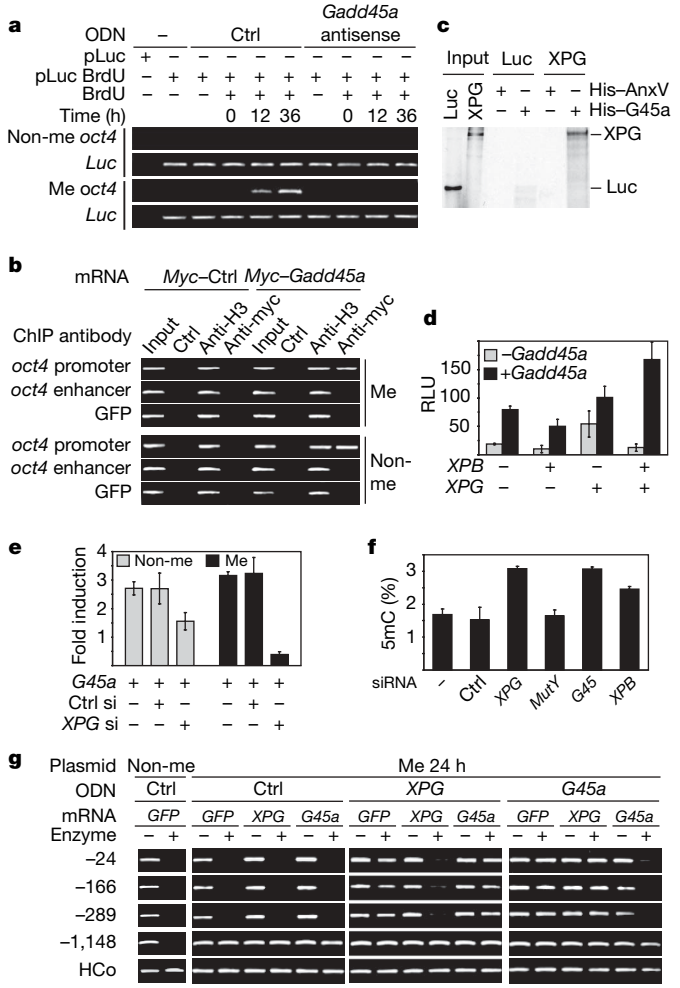


Figure 4 | Demethylation by *Gadd45a* involves DNA repair. **a**, DNA synthesis accompanying DNA demethylation requires Gadd45a. *Oct4* plasmid was injected with or without BrdU into *Xenopus* oocytes and recovered as described (Supplementary Fig. 4). To control for equal loading, *in vitro* BrdU-labelled luciferase plasmid (pLuc BrdU) was added after oocyte lysis. To control for immunoprecipitation specificity, unlabelled pLuc was added. PCR analysis of immunoprecipitated DNA using *oct4* or *Luc* specific primers was carried out. **b**, Gadd45a recruitment to the *oct4* promoter in *Xenopus* oocytes. Thirty six hours after mRNA injection, oocytes were injected into the nucleus with unmethylated or methylated *oct4* plasmid and harvested 6 h after DNA injection. ChIP analysis was performed using the indicated antibodies followed by PCR amplification of the *oct4* promoter or enhancer region or of *GFP*. Anti-H3, anti-histone-H3 antibody; Ctrl, control, rabbit serum. **c**, Gadd45a binds xeroderma pigmentosum complementation group G protein (XPG). SDS–polyacrylamide gel electophoresis and autoradiography of pull-down assays with recombinant His-tagged annexinV (His–AnxV, control) or Gadd45a (His–G45a) proteins and 35S-labelled *in vitro* translated Luciferase (Luc, control) or XPG. **d**, XPB and XPG synergize with Gadd45a. Luciferase reporter assays in HEK293 cells transiently transfected with methylated SV40 reporter. RLU, relative light units. Error bars, s.e.m. ($n = 4$). **e**, XPG is required for Gadd45a function. Luciferase reporter assays of HEK293 cells transiently transfected with non methylated or methylated SV40 reporter, *Gadd45a* and siRNAs. Error bars, s.e.m. ($n = 3$). **f**, *XPG* and *XPB* knockdown induces DNA hypermethylation. RKO cells were transfected with the indicated siRNAs. Capillary electrophoresis analysis of DNA 5mC as in Fig. 2d. Error bars, s.e.m. ($n = 4$). **g**, Active *oct4* demethylation in *Xenopus* oocytes requires XPG. PCR analysis of four *oct4* CpGs with (+) and without (−) methyl-sensitive restriction enzyme digestion as described in Fig. 3e, f. Oligodeoxyribonucleotides (ODN) against *VegT* (as control, Ctrl), *XPG* or *Gadd45a* were co-injected with *GFP*, human *XPG* or human *Gadd45a* mRNAs into the cytoplasm of stage VI oocytes as indicated. After 48 h, oocytes were injected into the nucleus with *oct4* plasmid and either harvested immediately (0 h, Supplementary Fig. 8) or 24 h after injection. Specificity of the ODN-mediated *XPG* knockdown is shown in Supplementary Fig. 7. HCo, *oct4* amplicon lacking *Hpa*II/*Hha*I sites as loading control.

XPG is an endonuclease responsible for the 3′ incision made during nucleotide excision repair[23]. We tested binding of mammalian XPG to Gadd45a by pull-down assay (Fig. 4c) and co-immunoprecipitation (Supplementary Fig. 5). In luciferase assays, *XPG* synergized with *Gadd45a* to activate methylated reporters (Fig. 4d). Interestingly, this synergy required the presence of the DNA helicase *XPB*, which associates with XPG (ref. 24). Conversely, siRNA knockdown of *XPG* (Supplementary Fig. 6) inhibited reporter activation by *Gadd45a* (Fig. 4e). Knockdown of either *XPG* or *XPB* (Supplementary Fig. 6) caused DNA hypermethylation (Fig. 4f). In *Xenopus* oocytes, ODN knockdown of *XPG* (Supplementary Figs 7–8) inhibited *oct4* demethylation at the proximal demethylation sites (Fig. 4g). This inhibition was rescued by co-injection of human *XPG* but not by *Gadd45a* mRNA. Likewise, *oct4* demethylation inhibited by *Gadd45a* knockdown was not rescued by *XPG* mRNA but by human *Gadd45a*, indicating that the two proteins cannot compensate for each other. We conclude that Gadd45a directly interacts with and requires XPG during active DNA demethylation.

The two main findings of this study are that *Gadd45a* is a key regulator of active DNA demethylation and that it acts by promoting DNA repair, thus linking both processes. Our data argue for a model whereby Gadd45a is targeted to specific sites of demethylation and recruits the DNA repair machinery through XPG. Methylated cytosines are then excised and replaced by unmethylated nucleotides. Although our data are consistent with a nucleotide excision repair mechanism, other repair mechanisms that also share components of the nucleotide excision repair machinery and involve nucleotide incorporation, such as long-patch base excision and mismatch repair, cannot be ruled out. Gadd45a interacts with DNA binding proteins such as histones[3] and nuclear hormone receptors[25], which may target Gadd45a to specific genomic sites.

Gadd45 proteins have been implicated in numerous biological processes and the finding that they relieve epigenetic gene silencing may account for some of these pleiotropic effects. For example, *Gadd45a* mutant mice show genomic instability and enhanced tumorigenesis[4], two hallmarks of human cancers that have been linked to altered DNA methylation patterns[1,26]. A possibility is that reduced *Gadd45* may lead to hypermethylation and inactivation of tumour suppressor genes such as *MLH1* (Fig. 3a, b). Finally, reprogramming by nuclear transfer and by exposure of permeabilized cells to *Xenopus* egg extracts[14,27] involves extensive DNA demethylation and *Gadd45* is likely to have a critical role in this process.

METHODS

DNA 'demethylase' screen. An arrayed *Xenopus tropicalis* embryonic cDNA library was screened by co-transfection of pools of 96 cDNAs[28] with 5 ng M.SssI-methylated SV40-luciferase reporter plasmid. Pools were scored as positive if they reactivated the methylation-silenced reporter. Positive pools were further selected by testing their ability to activate three non-related reporters. Individual cDNAs were isolated by sib selection of the arrayed, positive cDNA pools. *Gadd45a* was independently isolated three times after screening 1,044 pools, and the cDNA was subcloned into pCS2+.

Oocyte microinjection and methyl-sensitive PCR. Oocytes were collected, cultured, injected and lysed as described in Supplementary Methods. Recovered plasmid DNA was digested with *Pvu*II alone or *Pvu*II and either *Hpa*II or *Hha*I and PCR amplified as described[10]. Primers for the *Hpa*II site −289 and *Hha*I site −166 were: forward, 5′-CTGTAAGGACAGGCCGAGAG; reverse, 5′-GTGAAACCGTCCCTAGGTGA. Primers for the *Hpa*II site −24: forward, 5′-CCAAGGCAGGGGTGAGAG; reverse: 5′-CTCGGGGTGCCCACCTTCC. Control primers (HCo) were: forward, 5′-CAGCACTTCTCTGGGGTCTC; reverse, 5′-GCTGAGTGGGCTGTAAGGAC. −1,148 primers were as described[10]. Care was taken to ascertain that PCR amplification took place under non-saturating conditions.

BrdU Incorporation. Twenty injected oocytes were lysed as described above with some modifications. Two hundred pg of control DNA (pGL3-Pro Luciferase (Promega); unlabelled or *in vitro* BrdU-labelled by nick-translation, BrdU-Luc) was added during lysis to monitor the immunoprecipitation. Twenty μg of sheared herring sperm DNA (shsDNA) were added as carrier for precipitation. The DNA pellet was dissolved in 450 μl TE containing 50 μg shsDNA.

Immunoprecipitation, extraction and precipitation were performed as described previously[29] with modifications; 1 μg of anti-BrdU antibody (Sigma) and 20 μg of goat anti-mouse IgG (Dianova) were used; 20 μg glycogen were added as carrier for precipitation after immunoprecipitation; the DNA pellet was dissolved in 1 μl TE per oocyte and used for PCR. Primers for Luciferase plasmid were: forward, 5′-TCAAAGAGGC GAACTGTGTG; reverse, 5′-TTTTCCGTCATCGTCTTTCC. PCR of BrdU-Luc served as a loading control for anti-BrdU-immunoprecipitation efficiency. Primers for *oct4*: HCo primers (see above).

Received 9 October; accepted 4 December 2006.
Published online 31 January 2007.

1. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16,** 6–21 (2002).
2. Kress, C., Thomassin, H. & Grange, T. Local DNA demethylation in vertebrates: how could it be performed and targeted? *FEBS Lett.* **494,** 135–140 (2001).
3. Carrier, F. *et al.* Gadd45, a p53-responsive stress protein, modifies DNA accessibility on damaged chromatin. *Mol. Cell. Biol.* **19,** 1673–1685 (1999).
4. Hollander, M. C. & Fornace, A. J. Jr. Genomic instability, centrosome amplification, cell cycle checkpoints and Gadd45a. *Oncogene* **21,** 6228–6233 (2002).
5. Simonsson, S. & Gurdon, J. DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. *Nature Cell Biol.* **6,** 984–990 (2004).
6. Escher, G. *et al.* Demethylation using the epigenetic modifier, 5-azacytidine, increases the efficiency of transient transfection of macrophages. *J. Lipid Res.* **46,** 356–365 (2005).
7. Shou, W. & Dunphy, W. G. Cell cycle control by *Xenopus* p28Kix1, a developmentally regulated inhibitor of cyclin-dependent kinases. *Mol. Biol. Cell* **7,** 457–469 (1996).
8. Yeom, Y. I. *et al.* Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. *Development* **122,** 881–894 (1996).
9. Hattori, N. *et al.* Epigenetic control of mouse *Oct-4* gene expression in embryonic stem cells and trophoblast stem cells. *J. Biol. Chem.* **279,** 17063–17069 (2004).
10. Gidekel, S. & Bergman, Y. A unique developmental pattern of *Oct-3/4* DNA methylation is controlled by a cis demodification element. *J. Biol. Chem.* **277,** 34521–34530 (2002).
11. Furukawa, T. *et al.* Densely methylated *MLH1* promoter correlates with decreased mRNA expression in sporadic colorectal cancers. *Genes Chromosom. Cancer* **35,** 1–10 (2002).
12. Li, Q., Ahuja, N., Burger, P. C. & Issa, J. P. Methylation and silencing of the *Thrombospondin-1* promoter in human cancer. *Oncogene* **18,** 3284–3289 (1999).
13. Fang, J. Y., Lu, J., Chen, Y. X. & Yang, L. Effects of DNA methylation on expression of tumor suppressor genes and proto-oncogenes in human colon cancer cell lines. *World J. Gastroenterol.* **9,** 1976–1980 (2003).
14. Jaenisch, R. *et al.* Nuclear cloning, epigenetic reprogramming, and cellular differentiation. *Cold Spring Harb. Symp. Quant. Biol.* **69,** 19–27 (2004).
15. Prives, C. & Foukal, D. Use of oligonucleotides for antisense experiments in *Xenopus laevis* oocytes. *Methods Cell Biol.* **36,** 185–209 (1991).
16. Rhee, I. *et al.* CpG methylation is maintained in human cancer cells lacking *DNMT1*. *Nature* **404,** 1003–1007 (2000).
17. Smith, M. L. *et al.* Interaction of the p53-regulated protein Gadd45 with proliferating cell nuclear antigen. *Science* **266,** 1376–1380 (1994).
18. Smith, M. L. *et al.* Antisense GADD45 expression results in decreased DNA repair and sensitizes cells to u.v.-irradiation or cisplatin. *Oncogene* **13,** 2255–2263 (1996).
19. Choi, Y. *et al.* DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. *Cell* **110,** 33–42 (2002).
20. Jost, J. P. *et al.* 5-Methylcytosine DNA glycosylase participates in the genome-wide loss of DNA methylation occurring during mouse myoblast differentiation. *Nucleic Acids Res.* **29,** 4452–4461 (2001).
21. Gong, Z. *et al. ROS1*, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase. *Cell* **111,** 803–814 (2002).
22. Weiss, A., Keshet, I., Razin, A. & Cedar, H. DNA demethylation *in vitro*: involvement of RNA. *Cell* **86,** 709–718 (1996).
23. Reardon, J. T. & Sancar, A. Nucleotide excision repair. *Prog. Nucleic Acid Res. Mol. Biol.* **79,** 183–235 (2005).
24. Iyer, N., Reagan, M. S., Wu, K. J., Canagarajah, B. & Friedberg, E. C. Interactions involving the human RNA polymerase II transcription/nucleotide excision repair complex TFIIH, the nucleotide excision repair protein XPG, and Cockayne syndrome group B (CSB) protein. *Biochemistry* **35,** 2157–2167 (1996).
25. Yi, Y. W. *et al.* Gadd45 family proteins are coactivators of nuclear hormone receptors. *Biochem. Biophys. Res. Commun.* **272,** 193–198 (2000).
26. Issa, J. P. CpG island methylator phenotype in cancer. *Nature Rev. Cancer* **4,** 988–993 (2004).
27. Hansis, C., Barreto, G., Maltry, N. & Niehrs, C. Nuclear reprogramming of human somatic cells by *Xenopus* egg extract requires BRG1. *Curr. Biol.* **14,** 1475–1480 (2004).
28. Davidson, G. *et al.* Casein kinase 1 γ couples Wnt receptor activation to cytoplasmic signal transduction. *Nature* **438,** 867–872 (2005).
29. Cimbora, D. M. *et al.* Long-distance control of origin choice and replication timing in the human β-globin locus are independent of the locus control region. *Mol. Cell. Biol.* **20,** 5581–5591 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Sitter for help in characterizing XPG; H. Clevers, N. Giese, H. Schöler, and J. Smith for reagents; and G. Davidson and H. Steinbeisser for helpful discussions. This work was funded by the Deutsche Forschungsgemeinschaft.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to C.N. (niehrs@dkfz.de).

naturejobs

THE CAREERS MAGAZINE FOR SCIENTISTS

In 1992, US presidential candidate Ross Perot declared that he was expecting to hear a "giant sucking sound" as jobs rushed out of the country because of new free-trade agreements. When it comes to the pharmaceutical industry, jobs may be disappearing but globalization doesn't seem to be the culprit. The reasons are, in fact, a mix of mergers and acquisitions, potential blockbuster drugs failing to deliver on their promise, patents on other blockbusters expiring, and a wave of lawsuits and bad publicity over adverse effects of drugs that had cleared clinical trials.

Four years after becoming the world's largest drug company, Pfizer last month announced plans to lay off 10,000 people (see *Nature* **445,** 466–467; 2007). The move is just the latest in a series of job losses within the sector. In 2004, Bayer axed 550 jobs; a year later, Merck revealed plans to trim its global workforce by 11%, cutting some 7,000 jobs. Last week, Eli Lilly completed its takeover of Seattle-based ICOS, a move that could see up to 700 people lose their jobs. And if industry-savvy blogs are anything to go by, there are more cuts to come for this sector.

Just a decade ago, the drug industry looked like a safe bet for scientists seeking long-term job security. Now the tables have turned, as fewer companies are delivering the blockbusters their size demands. So what should you do if you still want to work in the industry? The answer lies in checking out potential employers' product pipelines. See what drugs a company has advancing through clinical trials and judge how your background and training might help bring these products to market. Increasingly, jobs in drug development are at 'big biotechnology' firms that have several products on the market, and the potential for more on the way. These companies tend to be staffed more with researchers than marketers or managers. Other opportunities exist at smaller firms, especially ones working in new areas beyond therapeutics, such as diagnostics and delivery. It's still not clear whether these opportunities will make up for the massive losses in pharmaceuticals. If they don't, we might be hearing not only a sucking sound but also a bubble bursting.

Paul Smaglik, *Naturejobs* editor

CONTACTS
Editor: Paul Smaglik
Assistant Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street,
London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
Business Development Manager:
Amelie Pequignot (4974)
e-mail: a.pequignot@nature.com
***Nature*events:**
Claudia Paulsen Young
(+44 (0) 20 7014 4015)
e-mail: c.paulsenyoung@nature.com

France/Switzerland/Belgium:
Muriel Lestringuez (4994)
UK/Ireland/Italy/RoW:
Nils Moeller (4953)
Scandinavia/Spain/Portugal:
Evelina Rubio-Morgan (4973)
Germany/Austria/The Netherlands:
Reya Silao (4970)
Online Job Postings:
Matthew Ward (+44 (0) 20 7014 4059)

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com
***Naturejobs* web development:**
Tom Hancock
***Naturejobs* online production:**
Catherine Alexander

US Head Office, New York
75 Varick Street,
9th Floor,
New York,
NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

Japan Head Office, Tokyo
Chiyoda Building,
2-37 Ichigayatamachi,
Shinjuku-ku,
Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe
e-mail: a.watanabe@natureasia.com

TAXI-CAB TEACHING

Short-term appointments are on the rise for teachers at colleges and universities around the world. Are these 'contingent' staff being taken for a ride? **Heidi Ledford** reports.

When Don Siegel went to graduate school at Rutgers University in New Jersey, he had his eye on the prize: a tenure-track professorship at a research university. It wasn't until years later, knee-deep in his research on protein structure and with a few semesters of teaching experience tucked under his belt, that he realized his priorities had shifted. So when a temporary, part-time position for a chemistry teacher came up, he grabbed it.

That was 1996. More than ten years later, Siegel is still at Rutgers, still teaching chemistry and still on a short-term contract. As a part-timer, Siegel bounced from assignment to assignment, without benefits such as health insurance or a retirement plan, until a year ago, when he switched to a short-term, full-time position managing the school's introductory chemistry course.

"It's been very frustrating and very satisfying at the same time," says Siegel. "I've had a dramatic impact on quite a few students. But in comparison to tenure-track people who've been here as long as I have, I'm making a fraction of what they're making."

Freeway flyers

Siegel is part of an expanding club of non-tenure-track, or 'contingent', teachers at universities and colleges around the world. Its members go by many names, depending on the institution and the country where they work. Officially, they may be called casual staff, adjunct professor, fixed-term worker or lecturer (which, confusingly, is a standard term for normal faculty posts in Britain). Unofficially, they're 'gypsy scholars', 'freeway flyers' or 'taxi-cab professors'.

Whatever you choose to call them, their numbers are growing. In Britain, 18% of the 2003 teaching and academic staff were short-term faculty members, up from 12% in 1995, according to the Association of University Teachers (now part of the University and College Union). In the United States, part-time and non-tenure-track full-time staff comprised 43% of the total faculty workforce in 1975 (see graph). By 2003, this had risen to 65%. Similar trends can be seen in Latin America and Asia: more than 80% of faculties in Argentina and Mexico have part-time appointments, says Philip Altbach, director of the Center for International Higher Education at Boston College in Massachusetts. As the ranks of the contingent teaching force swell, many wonder how the trend will change the face of higher education.

The driving force is economics. Budget-conscious institutions can save money by stitching together a patchwork of part-time workers. Such staff are often paid at a lower rate, and few institutions provide them with benefits. Jack Schuster, a professor of education and public policy at Claremont Graduate University in California, says using a part-timer for a course costs a fraction of what universities would pay a full-time tenured faculty member for the same teaching load.

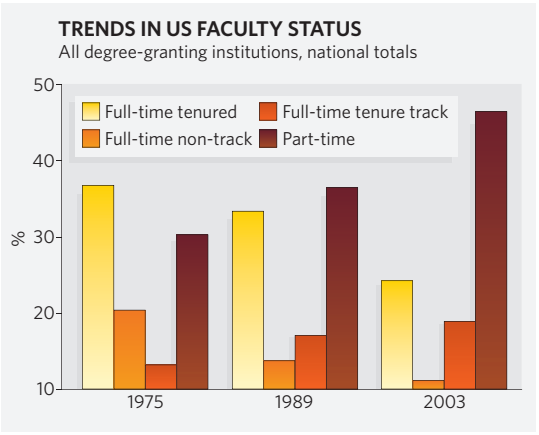
The economic advantage of hiring a full-time, non-tenure-track faculty member is not so clear. Although typically on short-term contracts, these teachers are often paid about as much as an entry-level tenure-track professor and typically (although not always) receive at least partial benefits. In this case, says Schuster, the institution is creating a dedicated teaching staff whose attention is not divided between research and teaching. And short-term contracts mean that the department can quickly adjust course offerings to fit the changing interests of the student body — an appealing option.

"If there's student interest for a particular kind of course, they hire the faculty," says Schuster. "And if that interest evaporates over the next several years, they are free to say 'We don't need you any longer.'" But, he adds, it remains to be seen whether this drive towards hiring non-tenure-track faculty members has improved teaching. Contingent faculty staff sometimes lack basic campus teaching resources such as offices, computers or even access to photocopiers. And part-timers who must rush from campus to campus have little time for meetings with students.

Siegel says short-term appointments give him little control over the courses he teaches. Tenured faculty members can sign up for classes a year or two ahead, but Siegel gets little warning of what his schedule will be. "There's no structure in place," says Siegel. "I teach where they tell me to teach. The only way I can be innovative is by going out on a limb and hoping that whoever follows in my footsteps will keep it going."

Sally Hunt is unconvinced by new work conditions.

A. WIARD



SOURCE: US DEPT EDUCATION, IPEDS FALL STAFF SURVEY

John Curtis, director of research and public policy at the American Association of University Professors in Washington DC, says that short-term appointments can also make contingent faculty members wary of taking risks. "At some point very soon they'll be looking for another job," he says. "For that reason, they tend to be in a position where they don't feel they can do anything controversial or risk offending students by challenging them."

Even if the teaching goes well, there is no guarantee that accepting a contingent position will lead to a tenure-track job down the road, says Monica Jacobe, a part-time instructor who shuttles back and forth between part-time teaching assignments at the University of Maryland in College Park and the Catholic University in Washington DC. Accepting a non-tenure-track post can be a good way to gain teaching experience, she says, but too much time there can become a liability. "Once you spend too many years — and that term is undefined — they start to think that there's something wrong and there's a reason that this person hasn't secured a full-time tenure-track job," says Jacobe, who is also a research fellow at the American Association of University Professors.

Out on a limb

Andrew Greene, a PhD adviser at the University of California, Berkeley, agrees that contingent faculty members can become typecast. Spending more than two or three years in fixed-term positions can be a liability in the sciences, he says, because research opportunities are limited. He adds that even small liberal arts colleges want faculty candidates to have a strong, current research background. "These days, places that 20 years ago had no great ambitions to have their faculty publish have ratcheted up their expectations," says Greene.

But it's difficult for contingent faculty members to win extramural research grants, and internal travel and professional-development grants are often limited to permanent staff. Meanwhile, a heavy teaching load leaves little time for research and publishing. Full-time contingent faculty members in the United States teach an average of 50% more than their tenure-track peers, according to Ernst Benjamin, a consultant to the American Association of University Professors.

Philip Altbach sees worrying trends in faculty employment.

Steve Wilson, a maths instructor at Sonoma State University in Rohnert Park, California, has applied for 13 tenure-track positions in the past 15 years, many of them at Sonoma State. He didn't get any. "It was very upsetting," he says. "One of the reasons they were able to hire people at Sonoma State is because I had put a lot of energy into increasing enrolment in maths classes."

Meanwhile, Siegel has also unsuccessfully applied for tenure-track positions. He is not sure if his extended time as a contingent faculty member at Rutgers has helped or hurt his chances. "A lot of schools are more comfortable taking a contingent faculty member than, for example, someone fresh out of industry," Siegel says. "But on the other hand, research is sexy."

Greene recommends limiting temporary positions to only a few years, and trying to stay in contact with the research community by attending meetings and even continuing a little postdoctoral research on the side if possible. But all of these are just temporary fixes for the bigger problem — a growing pool of PhDs and a shrinking pool of tenure-track positions. "Right now, it's a buyers' market for academic labour," says Wilson.

Hired and backfired

Unfortunately, some attempts to better the plight of temporary academic workers have backfired. The state of New Jersey, for example, declared that full-time, non-tenure-track faculty members should be automatically switched to the tenure track after four years. Regulations such as this, although well-meaning, have caused some universities to fire short-term staff. "The general practice is that the contingent faculty stay for a limited period, and then get booted out," says Altbach.

Sally Hunt, joint general-secretary of the University and College Union, a UK academic trade union, says she is worried that recent European Union directives to improve conditions for fixed-term workers may have similar results. That directive required member states to specify a maximum duration for temporary contracts before an employee becomes permanent. In response to the rule, Britain mandated that fixed-term workers who have been at an institution for four years automatically switch to a permanent contract unless the employer objects. But Hunt says that the conditions under which an employer can object are undefined, allowing institutions to fire fixed-term workers before the end of four years without sufficient cause.

Nevertheless, at some institutions, contingent faculty members have gained ground. Wilson says that the California State University system now offers three-year contracts to contingent faculty members who have worked in the system for more than six years. And David Robinson of the Canadian Association of University Teachers says that some Canadian universities now have seniority clauses that give contingent faculty members advance notice of new tenure-track job openings.

Jacobe says that changes such as these are crucial, given the growing use of contingent faculty members. "If this is going to be the reality of academia," she says, "then we need to make it a workable reality." Nevertheless, she wants to continue teaching. "I think we all do this because we like teaching," says Jacobe. "But liking it doesn't mitigate everything that's wrong with the system." ■

Heidi Ledford is a science writer in Berkeley, California.

NATURE|Vol 445|8 February 2007

MOVERS

Daniel Zajfman, president, Weizmann Institute of Science, Rehovot, Israel

2005–06: Director, Max Planck Institute of Nuclear Physics, Heidelberg, Germany
1997–present: Associate then full professor, Department of Particle Physics, Weizmann Institute of Science, Rehovot, Israel
1991–97: Senior scientist, Department of Particle Physics, Weizmann Institute of Science

Daniel Zajfman had a simple motivation when he enrolled in physics in 1979 at the Technion, Israel Institute of Technology: a desire to understand nature. He has followed that ambition throughout an accomplished career.

In December, Zajfman became the tenth president of Israel's Weizmann Institute of Science in Rehovot. He lauds the frequent interactions between theoretical and experimental physics at the institute — one of its trademarks. Now he hopes to encourage interactions among other fields as well, such as biology and biochemistry. Zajfman says he would consider regrouping researchers into interdisciplinary departments.

Born in Belgium in 1959, Zajfman moved to Israel at the age of 20. In 1989 he received his PhD in atomic physics from the Technion, then spent two years as a postdoctoral fellow at the Argonne National Laboratory near Chicago. When he returned to Israel in 1991, he joined Weizmann's department of particle physics as a senior scientist.

Zajfman also has strong ties with Germany. Since 2001 he has been an external member of the Max Planck Institute of Nuclear Physics in Heidelberg, where he was appointed a director in 2005. He found that science was an excellent means of overcoming political tensions among Israelis and Germans. "The diplomatic story between Israel and Germany started with scientists," he says.

The Weizmann Institute has no problem recruiting graduate students, says Zajfman, although political instability has made it more difficult to attract foreign postdocs. He believes there is a widespread misconception about life in Israel. "Some people think it's a third-world country," he says. The political situation does, however, make life more complicated. "You organize an international conference with 200 scientists in Eilat," says Zajfman, "then a bomb explodes in Jerusalem, which is 300 kilometres away, and the whole conference is cancelled."

Zajfman will face other challenges, says Andreas Wolf of the Max Planck Institute of Nuclear Physics. "He undertakes the task of ensuring further funding of the Weizmann Institute," says Wolf, noting that private financing plays a big role there. One of Zajfman's biggest challenges will be communicating the ideas behind the science.

At 47, Zajfman is the youngest president in Weizmann's 60-year history. He laments that he'll have less time for science and for his students, but he won't stop his research altogether. Exploring nature remains his primary passion. ■

Nora Eichinger

NETWORKS & SUPPORT

Seeking a PhD abroad

Although many budding scientists get international experience as part of their postdoc, acquiring international exposure during one's PhD is less common. But it can be just as rewarding.

I'm in the final year of my PhD course in experimental condensed-matter physics, studying magnetism using a technique called neutron scattering. The PhD programme is based at Loughborough University, UK, but a big part of my experimental work is at the Max Planck Institute for Solid State Research in Stuttgart, Germany. Coming from England just after my bachelor's degree in physics was a big step and somewhat risky, both personally and professionally. I wasn't sure it would help my career or suit my personal style. I was faced with a number of questions. Should I stay for further PhD study at my home institution? What topic will help me pave the desired career path? If I do go abroad, where to?

My supervisors, one in England and one in Germany, gave me the freedom to pick research topics and carry out my research at another institute. I was excited about going to the Max Planck Institute, which has people from more than 35 different countries at the PhD and postdoc level.

Once the long road of the PhD is complete, I hope to have the confidence to compete in an increasingly multinational field in order to secure postdoc positions. Although moving abroad means reduced contact with my home university and research community, attending conferences on neutron scattering in Britain has helped me network and develop a sense of the community.

It took some time to settle down, find an apartment and get used to the local life, surroundings and a different language. In the lab, a simple task such as ordering helium for my experiments is complicated by the language barrier. In a few cases, this meant accidentally ordering extra tanks of the stuff.

Still, with persistence, it's been a rewarding and enriching experience both personally and professionally. The Max Planck Institute provides not only a stimulating environment, but ample funding that has helped boost the output of my research considerably. I have access to equipment and resources hard to find at many institutions.

Thus far, I have no regrets. Doing a PhD abroad has improved my knowledge of physics and clarified my career aspirations. ■

Michael Banks is a PhD student at the Max Planck Institute for Solid State Research in Stuttgart, Germany.

POSTDOC JOURNAL

Making a difference

So, what's next? So simple a sentence, and yet perhaps the most terrifying question in the English language for students nearing the end of their doctorate. Should I do a postdoc? The guilt sets in. Didn't we all tell our graduate-school admissions committees that we were planning careers in academia? Aren't we newly minted PhDs now supposed to be aiming for that coveted assistant professorship? There, but for the long and treacherous postdoctoral road, I go.

And yet, against all my best intentions, I'm on that very road. If you'd told me two years ago, in the midst of my mid-PhD slump, that I'd be continuing in research through a postdoc, I would have laughed, and then cried, and then quit and returned to my native Australia. What went wrong?

I'll tell you — it's the allure of making a difference. I hoped to do so with my doctorate on cardiac arrhythmias, but I foresee only incremental improvements to existing therapies. Now I'm switching fields completely, researching nutrition and human body composition regulation, partly because addressing malnutrition and obesity promises to mitigate global public-health problems.

So, I'm giving myself two years, three at the most, to determine whether I can indeed make a palpable difference. Will my published articles effect change? Will mixing with experts excite me? Will my work have public-policy implications? ■

Peter Jordan is a first-year visiting fellow at the National Institute of Diabetes & Digestive & Kidney Diseases in Bethesda, Maryland.

The inside track from academia and industry

Time for a change

Universities, businesses, students and employees all need to rethink they way they approach training.

Michael Alvarez

Advancing science through academic research isn't inherently at cross-purposes with the scientific trainee's career progression — but sometimes it seems that way. There is an inherent tension between the two, because universities need graduate and postdoctoral researchers to keep labs productive, while the graduate students and postdocs require training, independence and permanent positions. For institutions to align these endeavours successfully, they need to consider greater career support for scientific research trainees.

A broad, long-standing practice in business called 'change management' provides frameworks and methods for helping organizations move from an existing model of operation to a more adaptive one. Science could be well served by borrowing one of its processes called 'merging agendas' to bring together the goals of early-stage scientists and of the universities where they train.

Put simply, this process involves identifying areas where the interests of different parties overlap, and acknowledging that interdependencies exist. This is an important part of any successful strategy, as resistance to change from those adopting a new system or process typically decreases as the awareness of mutual dependence, goals and benefits increases. In addition to moving the parties beyond resistance, the process can actually help engender enthusiasm, energy and excitement for change.

It is clear that different university environments are unique, and there is no single model or approach to enhancing career services that can simply be copied from one campus and taken to the next. Nevertheless, there is common ground in the challenges that universities and their trainees face, in light of their shared desire for high-quality research and good career progression. Despite the challenges involved in making such changes, there is some encouraging evidence of a trend towards improving career support for scientific researchers within certain academic environments.

In my experience establishing the medical and life-sciences career centre at the University of California, San Francisco, and now at Stanford University, I have seen a variety of approaches to enhancing the training experience to include more adequate career preparation. From grass-roots student-led initiatives, to fully fledged university commitments to building comprehensive career centres — and others in between — the efforts can yield widely varying degrees of success.

On the basis of input and observations gathered from administrators, deans, career services professionals, trainees and scientific journals that I have consulted along the way, I can say there are some noticeable characteristics (see 'Who should be doing what', below) that bode well for bringing the agendas of trainees and universities together.

As within any campus community, there will be a range of mindsets, biases and/or priorities represented; some agendas will be revealed, some hidden. To successfully bring research and training into greater alignment, the participating groups must reconsider their roles and commit to making the necessary changes. Once people accept that assuming these responsibilities is non-threatening and in fact healthy, performance can be mobilized and greater results achieved at both the individual and the organizational level.

Trainees must think more about how to help themselves and take ownership for their careers, although university administrators and faculty members must be willing to see the promotion of the trainees' careers as one of their significant duties or obligations. To advance the research cause of an academic institution without attention to the career development of the individual trainees is a formula that may have worked in the past. But, given the needs and challenges of the global labour market today, it seems very much less tenable.

By looking at the respective agendas through a common lens, the two endeavours can be seen as parallel and beneficial to one another. Although the process of change is a slow one, it is possible and even inevitable. The people and institutions that embrace the opportunity to blend their agendas are likely to be at the leading edge, both in their research and in their careers. ■

Michael Alvarez is director of Stanford University's School of Medicine Career Center, Stanford, California.

WHO SHOULD BE DOING WHAT

University/institution obligations and responsibilities

- Departments and faculty members must accept the need to change and reject the status quo.
- Schools and universities must show commitment to strengthen career resources.
- Qualified, professional career advisers should be available for confidential services.
- Deans and administrators must work together to address problems in the system.

Trainee obligations and responsibilities

- Structure formal steps to explore careers outside academic research.
- Do not assume that job-market and career matters will simply work out on their own.
- Remain active in career planning throughout the training experience.
- Understand and evaluate multiple career choices before making decisions.